SKILLS, PERSONAL CHARACTERISTICS, AND LABOR MARKET TRANSITIONS

Dissertation

submitted to the

Faculty of Business, Economics and Informatics

of the University of Zurich

to obtain the degree of

Doktor der Wirtschaftswissenschaften, Dr. oec.

(corresponds to Doctor of Philosophy, PhD)

presented by

Peter Höschler

from Germany

approved in February 2018 at the request of

Prof. Dr. Uschi Backes-Gellner

Prof. Dr. Carmit Segal

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 14.02.2018

The Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Acknowledgments

I begin by thanking my doctoral advisor, Professor Uschi Backes-Gellner, for her outstanding supervision, guidance, excellent knowledge, and constant support throughout my graduate studies. I am very grateful that she believed early in my capabilities and supported me in applying for this unique doctoral program. She has given me valuable advice and contributed a great deal to my growth both inside and outside of academic research.

I would like to express my gratitude to my doctoral co-advisor, Professor Carmit Segal, who supported me throughout my graduate studies. She contributed significantly to the improvement of my academic skills through her valuable comments and inspiring ideas. I gratefully acknowledge Professor Ofer Malamud for making my research stay at the University of Chicago possible. His time and valuable comments substantially improved my research. I am also grateful to Professor Edward Lazear and Professor Stefan Wolter for their helpful feedback during our meetings and seminars. For editorial help and guidance, my thanks go to Natalie Reid.

I gratefully acknowledge financial support from the Swiss National Science Foundation (SNSF) for my visiting scholarship at the University of Chicago and from the Swiss State Secretariat for Education, Research and Innovation (SERI) through its Leading House on the Economics of Education, Firm Behavior and Training Policies.

My thanks go to colleagues at the University of Zurich and the University of Chicago for constructive discussions and the provision of a collegial working atmosphere. In particular, I would like to thank Simone Balestra for his invaluable guidance and support along the way; Michael Ribers and Florian Schaffner for being enthusiastic fellows during the course work and over the years; Simon Janssen, Johannes Kunz, and Philippe Ruh for empirical guidance; and Christa Pálffy, Curdin Pfister, and Simone Tuor for their constant support. Thanks also go to Benjamin Boller, Lars Mehr, and Jason Blunier for their excellent research assistance.

I extremely appreciate the fantastic group of people—from both inside and outside of academia—whom I am pleased to call my friends. They have constantly empowered me over the years. I admire and value every one of them.

Finally, I am deeply grateful for the support and love of my family. I thank my grandparents for supporting me in starting this journey. I thank my sister Sandra for always being available for advice and encouragement. Finally, I wish to thank my parents, Daniela and Thomas, for all the support and motivation during all those years, making it possible for me to pursue this education. My parents' unconditional regard is the seed of my academic growth. I will always be grateful for their belief in me.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Character is higher than intellect.

Ralph Waldo Emerson

Nature; Adresses, and Lectures (1849)

That skills and personal characteristics are multidimensional is a common belief. Moreover, social scientists have long acknowledged the various facets by which individuals can be described. During the 20th century, this mere acknowledgment became a solid empirical finding, as psychologists started to provide a toolbox of measures for various personal characteristics. Thus both inside and outside of academia it is now well established that a variety of personal characteristics, including cognitive ability and non-cognitive skills, are important for success in life.[1]

Since the early 2000s, economists have shown increasing interest in non-cognitive skills for two major reasons: the labor market relevance of these skills and their malleability. This movement was triggered by the provocative book, *Bell Curve*, by psychologist

[1] I use the term *non-cognitive skills* to emphasize the malleability of these skills, and the term *cognitive ability* to emphasize the higher inherence of cognition. *Personal characteristics* refers to a bundle of skills, abilities, and traits, which includes cognitive ability and non-cognitive skills. *Personality traits* are a subset of non-cognitive skills.

Richard Herrnstein and political scientist Charles Murray (Herrnstein & Murray, 1994). They paint a deterministic picture of western society by showing, first, that cognitive ability is uniquely important and, second, by assuming that it is primarily inherited and thus not subject to interventions. In his response, Heckman (1995) argues against both of these statements and subsequently initiates a literature in economics that investigates the importance of multiple skills for success in life.

Economists have clearly shown that besides cognitive ability, non-cognitive skills are also highly valued in the labor market (for an overview, see Almlund, Duckworth, Heckman, & Kautz, 2011; Heckman & Kautz, 2012). A large body of literature shows the positive effect of non-cognitive skills—including personality traits, self-esteem, and locus of control—on earnings (e.g., Goldsmith, Veum, & Darity, 1997; Heckman, Stixrud, & Urzua, 2006; Heineck & Anger, 2010; G. Mueller & Plug, 2006; Murnane, Willett, Braatz, & Duhaldeborde, 2001; Nyhus & Pons, 2005). Heckman, Pinto, and Savelyev (2013) find that the returns to early childhood education are driven mainly by the program's effect on the development of non-cognitive skills. Moreover, by revealing that achievement tests such as the General Educational Development testing exam (GED) are no equal substitute for high school graduation as they do not measure important skills, Heckman, Humphries, and Kautz (2014) reinforce the value of non-cognitive skills for succeeding in life. Another very recent strand of literature emphasizes the value of social skills (e.g., Deming, 2017). The bulk of this literature shows the value of non-cognitive skills in the labor market.

At the same time, non-cognitive skills are indeed highly malleable (for an overview, see Cunha, Heckman, Lochner, & Masterov, 2006). Early childhood education, particularly the Perry Preschool Program, is the major intervention studied in this regard (e.g., see Heckman et al., 2013). Nevertheless, critical periods exist for the development of both cognitive ability and non-cognitive skills, periods that may exist even later in life for the development of non-cognitive skills (Cunha & Heckman, 2008). Indeed, as this malleability allows economists to draw action-oriented policy implications, it is key for their interest in extensively studying these skills.[2]

This dissertation adds to both strands of the economic literature on non-cognitive skills: that on labor market relevance and that on malleability. The first two projects investigate the malleability of non-cognitive skills during adolescence and adulthood. The third project connects the two strands by investigating the labor market relevance of both the initial levels of non-cognitive skills and their changes over time. It also provides a comparison between the predictive power of cognitive ability and that of non-cognitive skills for an important labor market outcome, i.e., receiving a job offer. All three projects investigate important non-cognitive skills and thereby share a tight labor market closeness as they investigate effects while the studied individuals are either transiting to the labor market or already working in it. In sum, the guiding theme of this dissertation is the investigation of labor-market relevant non-cognitive skills.

After this introductory chapter, chapters two and three investigate the malleability of non-cognitive skills. In chapter two, I investigate the development of self-esteem, a highly labor market-relevant non-cognitive skill (e.g., Drago, 2011; Heckman et al., 2006), during adulthood. Using the U.S. National Longitudinal Survey of Youth 1979, I investigate the long-term relationship between failure in general education and self-esteem. While the results for the relationship between education and self-esteem are mixed (De Araujo & Lagos, 2013; Heckman, Humphries, Veramendi, & Urzua, 2014), I investigate the role of unfulfilled educational aspirations in the development of self-esteem. Moreover, I explicitly add to the literature by investigating the benefits of community colleges in great detail. In my analysis, I am able to investigate long-term effects from when the former students are in their 40s. Moreover, I link an explicit life event, dropping out of college, to the development of non-cognitive skills in a quasi-interventional way.

[2]If all skills are set in stone at birth, all studies investigating labor market returns to these skills would merely describe a form of discrimination. To the contrary, studies show, for example, that gender discrimination in wages can partly be explained by differences in non-cognitive skills (e.g., Heineck & Anger, 2010; G. Mueller & Plug, 2006), thereby allowing clear policy implications which target the formation of non-cognitive skills.

My results imply that a more diverse education system with multiple paths might be beneficial for the formation of self-esteem. As vocational education and training might be such a path, I investigate the development of non-cognitive skills during this type of education in the next chapter.

In chapter three, I use a unique, self-collected[3] longitudinal data set—the Leading House Apprenticeship Panel—to study the development of Grit (i.e., the perseverance and passion for long-term goals, see Duckworth, 2016) and the Big Five personality traits (conscientiousness, extraversion, agreeableness, openness, and emotional stability) while adolescents are enrolled in apprenticeship training. Although most of the literature suggests that non-cognitive skills are malleable during childhood (for an overview, see Cunha et al., 2006), less is known about the formation and malleability of these skills during adolescence. Moreover, this study is the first to investigate the development of Grit over a period spanning six years compared to the so far documented development over one year (Duckworth & Quinn, 2009).

The results of both projects underscore the malleability of non-cognitive skills after childhood. In chapter two, I find that unfulfilled educational aspirations have long-term negative effects on the development of self-esteem. To distinguish college dropouts with unfulfilled aspirations from graduates with fulfilled aspirations, I develop a classification of education combining the highest type of college enrolled in (aspiration) and the highest degree obtained (realization of aspiration). Using data spanning three decades, I find that four-year college dropouts, when compared with graduates, have permanently lower self-esteem, whether the dropouts obtained an associate's degree or not. However, associate degree holders who had never enrolled in a four-year college do not experience this long-term negative effect. Therefore, finishing the highest type of college in which the student ever enrolled—that is, fulfilling his or her educational aspirations—is critical for the formation of self-esteem. To achieve this fulfillment, students need to have realistic

[3]The data collection was initiated in 2009 as part of Donata Bessey's dissertation (Bessey, 2010) and continued as part of Yvonne Oswald's dissertation (Oswald, 2013) until 2013. As part of this dissertation, I conducted the final follow-up surveys in 2014 and 2015.

aspirations.

In chapter three, for Grit, I find significant within-person mean-level increases of about .5 standard deviation units over a six-year period. For the Big Five personality traits, I estimate similar increases for conscientiousness, agreeableness, and emotional stability. I show that these changes are heterogeneous and robust to reasonable levels of measurement error. Moreover, to explain the heterogeneity in skill changes, I use an educational production framework. This analysis reveals that workplace-based factors—e.g., the fulfilment of job expectations, the satisfaction with prospects of income, or the perceived workload—as opposed to educational resources, play a crucial role in explaining the heterogeneity in the development of non-cognitive skills. In sum, my results show that non-cognitive skills develop during vocational education and training but that this development is heterogeneous over individuals. Further, the heterogeneous development is partly due to differences in workplace-based factors and not primarily due to differences in educational resources.

Chapter four follows up on these results by investigating the labor market relevance after apprenticeship training of both, the initial levels of non-cognitive skills and their changes over time. Therefore, I examine the relative importance of different personal characteristics for hiring decisions. To asses this relative importance, I investigate firms' revealed preferences during the retention process at the end of training. I use the unique data set introduced in chapter three, data that provides me with high-quality, well-established measures of intelligence, grades, economic preferences, Grit, and the Big Five personality traits. Apprenticeship training with its duration of three to four years enables firms to observe potential workers over a long screening period. When the training ends, firms can decide to make job offers to certain apprentices, thereby revealing their preferences about workers' personal characteristics. I connect real-world job offers and workers' personal characteristics, both of which are usually unobserved by researchers. To investigate the relative importance of various personal characteristics for the likelihood to receive a job offer, I conduct a comparison between the predictive power of various personal characteristics, including cognitive ability and non-cognitive skills. This project answers two major questions: Which personal characteristics do employers value? Are non-cognitive skills related to job offers, i.e., to employment? The second question investigates an important avenue through which non-cognitive skills may affect labor market outcomes such as wages.

The results in chapter four underscore the importance of non-cognitive skills in the labor market. I find that grades and non-cognitive skills are important for receiving a job offer, with the Big Five personality traits by far the most important predictor. For Grit, I find that its development during apprenticeship training, rather than its initial level, is the main predictor. I find no effects for either intelligence or economic preferences. In sum, the results of this chapter add to the literature on the labor market-relevance of non-cognitive skills. As my results show that firms use non-cognitive skills in their hiring decisions, they therefore provide evidence that non-cognitive skills are related to employment which subsequently affects other labor market outcomes. Therefore, they show a mechanism through which non-cognitive skills affect several labor market outcomes.

Taken together, the findings in chapters two through four clearly indicate three important implications. First, the results in chapter two point to the value of a diverse education system with multiple paths that have different features. Beyond general education, vocational education and training might provide an additional path. Second, the results in chapter three and their interpretation in the light of the literature on the stability of non-cognitive skills during general education (Elkins, Kassenboehmer, & Schurer, 2016) indicate that vocational education and training might foster the development of non-cognitive skills. Third, the results in chapter four reveal the importance of screening periods—an implicit element of apprenticeship training—as a tool for employers to select potential entry-level workers. In such a period non-cognitive skills become observable and might even further develop which in turn affects jobs offers afterwards. In sum, the results in these three chapters underscore the importance of vocational education and training.

Chapter five synthesizes the key findings of chapters two through four and concludes. It also discusses further research questions and policy implications that follow from the findings of this dissertation.

Chapter 2

Failing in General Education: College Dropout and Self-Esteem

Part of this chapter is a revised version of early parts of the working paper "Shooting for the Stars and Failing: College Dropout and Self-Esteem" by Hoeschler and Backes-Gellner (2014).

2.1 Introduction

That self-esteem is an important non-cognitive skill highly valued in the labor market is well known. Several economics studies show a significant positive, persistent, and partly causal effect of self-esteem—or an aggregate of self-esteem and other non-cognitive skills—on wages (Drago, 2011; Goldsmith et al., 1997; Heckman et al., 2006; Murnane et al., 2001; Waddell, 2006). These findings underscore the importance of learning how individuals maintain and develop self-esteem. Life events might affect self-esteem. According to Bénabou and Tirole (2002), rational individuals—at least to some extent—recall "bad news" to infer their self-beliefs. Thus some empirical studies have shown that adverse life events result in lower levels of self-esteem (e.g., see Goldsmith, Veum, & Darity, 1996, for the negative effects of joblessness on self-esteem, or Prevoo & ter Weel, 2015, for the negative effects of family disruptions on self-esteem). To better understand the effect of adverse events related to education, we study the impact of dropping out of college, one very concrete and widespread life event, on self-esteem.

We define dropouts as individuals who failed to attain a degree in the highest educational program they aspired. We investigate the relationship between these unfulfilled educational aspirations and self-esteem. For this purpose, we develop an educational classification, which enables us to separate college dropouts and graduates. In this classification we use college enrollment as a measure for revealed educational aspirations. Empirical research on whether and through what mechanisms education affects self-esteem (or the vice versa) does not take into account educational aspirations. For example, studies using grade point average (Himmler & Koenig, 2012) or completed years of schooling (De Araujo & Lagos, 2013) as the educational variable, view self-esteem primarily as an input in the education production process and report effects of self-esteem on academic achievement. However, a classification of education based on completed years of schooling does not provide clear information on college enrollment, i.e., aspirations, and on degree obtainment, i.e., realization of aspirations, information critical to investigate the role of educational aspirations.[1]

Heckman, Humphries, Veramendi, and Urzua (2014) use degree obtainment and college enrollment to classify final schooling outcomes. With their classification they are able to show that college enrollment (versus no college enrollment, conditional on the individual being a high school graduate) and earning a four-year degree (versus some college)

[1]First, the completed years of schooling differ from the typical years of schooling attributed to a degree for a substantial share of college degree holders (Jaeger & Page, 1996). Thus years of schooling is not a clear measure for differentiating college dropouts from graduates. Second, individuals who enroll in college but never finish their first year of studies—a group that clearly consists of college dropouts—have completed 12 years of schooling and therefore, have completed the same years of schooling as high school graduates who never have enrolled in college (Park, 1996).

causally improve self-esteem. However, they do not fully take into account educational aspirations as they only focus on four-year college completion. Given that educational decisions in a system with two- and four-year colleges can be more complex, a classification based only on four-year college degrees cannot fully capture college completion and non-completion patterns, particularly, as college-enrolled individuals without a four-year college degree could still have graduated from a two-year college. Therefore, more detailed classifications of college education are necessary for estimating the relationship between unfulfilled educational aspirations and self-esteem.

Such detailed classifications of college education are well established in other study fields, and they have shown the fruitfulness of clearly distinguishing between college graduates and dropouts of different types of colleges. For example, more detailed classifications of college education are used for estimating the heterogeneous labor-market returns to enrollment in and graduation from different types of colleges. In particular, this literature shows substantial returns to graduating from a two-year college with an associate's degree (Belfield & Bailey, 2011; Jepsen, Troske, & Coomes, 2014; Kane & Rouse, 1995; Liu, Belfield, & Trimble, 2015; Zeidenberg, Scott, & Belfield, 2015). These returns actually appear to outperform the returns of merely attending a four-year college without receiving any degree (Kane & Rouse, 1995). Thus these findings indicate the importance of applying a detailed classification of final schooling outcomes, a classification that distinguishes between two- and four-year colleges, for studying the consequences for labor market or other outcomes.

In this paper, we develop an educational classification that provides us with a clear-cut definition of college dropouts, which enables us to investigate the relationship between unfulfilled educational aspirations and self-esteem. Taking into account the complexity of post-secondary educational decisions, we use educational paths, i.e. we combine college attendance as a measure for educational aspirations and degree obtainment as a measure for educational attainment (realization), to distinguish among different final college outcomes. We classify individuals by their combination of the highest type of college ever attended (two- or four-year college) and the highest degree ever received (high school degree, associate's degree, or bachelor's degree). In this framework, we define individuals as "dropouts" whenever the highest degree they received (attainment) is lower than the degree that the highest type of college they ever attended would usually grant (aspiration). Therefore, our classification of education gives us a clear definition of dropouts, one that enables us to investigate the relationship between dropping out of college, i.e. unfulfilled educational aspirations, and self-esteem.

We extract a sample of college-enrolled individuals from the U.S. National Longitudinal Survey of Youth 1979 (NLSY79), allowing us to conduct an investigation of the relationship between dropping out of college and self-esteem over a long period. By exploiting the panel structure of the NLSY79, we find that among individuals who enrolled at any point in a four-year college, dropping out—compared to graduating—results in significantly lower self-esteem. When considering all college-enrolled individuals, we find that two- and four-year college dropouts have significantly lower self-esteem than four-year college graduates. This finding also holds for four-year college dropouts who moved between two- and four-year colleges but received only an associate's degree. In contrast, two-year college graduates who never enrolled in a four-year college have no lower self-esteem. In sum, all dropouts—i.e. all students with unfulfilled educational aspirations—have lower levels of self-esteem. Given the labor market relevance of self-esteem, we also show that these differences in self-esteem can be linked to wage differences.

Taken together, our results imply that students should aim at forming realistic educational aspirations and enrolling only in colleges from which they can reasonably expect to graduate. Policy makers may support this process by increasing permeability in the education system which would allow individuals to progressively gain educational credentials, and, thereby, self-esteem. In such a system, students could start first with acquiring lower levels of education, and then moving step-by-step up to their optimal level of education.

2.2 Development of Self-Esteem during Education

It is well-established in the psychology literature that self-esteem develops over the life-cycle (e.g., see Orth, Robins, & Widaman, 2012; Robins, Trzesniewski, Tracy, Gosling, & Potter, 2002; Trzesniewski, Donnellan, & Robins, 2003). Moreover, several studies show that academic achievement influences the formation of self-esteem. Baumeister, Campbell, Krueger, and Vohs (2003) intensively survey the non-economics literature on the relationship between self-esteem and education. They conclude that high self-esteem is the result of school performance and that any existing causal relationship goes from academic achievement to self-esteem.

More recently, economists have also started to investigate the relationship between schooling and self-esteem. While we argue that differences between educational aspiration and attainment affect self-esteem, these studies focus on the effects of other educational variables. Heckman et al. (2006) show that discrete schooling outcomes causally increase self-esteem. For example, individuals with 13 or more completed years of schooling score higher on self-esteem than individuals who have just completed 12 years, that is, who merely finish high school. However, studies—using completed years of schooling as a continuous variable (De Araujo & Lagos, 2013) or grade point average (Himmler & Koenig, 2012) as the educational variables in instrumental variable estimations—find effects of self-esteem on education. These contradicting findings show the importance of the classification of education when investigating the relationship between education and self-esteem.

To investigate the effect of educational decisions on labor market, health, and social outcomes, Heckman, Humphries, Veramendi, and Urzua (2014) develop a sequential model of educational decisions, including four decision nodes: high school graduation, GED obtainment, college enrollment, and college graduation. For self-esteem (measured when the individuals are in their 40s), final outcomes of education do not causally increase self-esteem when Heckman, Humphries, Veramendi, and Urzua (2014) compare the final outcomes to the base group of high school dropouts. When they focus on a specific decision node in the sequential educational decision model, their estimates are more precise, and they find two significant results: First, enrollment in college, compared with not having enrolled in college after high school graduation, increases self-esteem; second, earning a four-year college degree compared with only enrolling in college without a four-year college degree (some college) increases self-esteem. Thus, focusing on educational decisions in detail, Heckman, Humphries, Veramendi, and Urzua (2014) show causality from college education to self-esteem. However, Heckman, Humphries, Veramendi, and Urzua (2014) do not investigate the role of educational aspirations. Therefore, we build on Heckman, Humphries, Veramendi, and Urzua (2014) and extend their classification to fully include educational aspirations.

Using a detailed classification of educational outcomes that takes different types of colleges into account and that enables us to differentiate between college dropouts and graduates from these different types of colleges, we investigate the effect of educational aspiration and attainment on self-esteem. We specifically target a group that Heckman, Humphries, Veramendi, and Urzua (2014) do not differentiate: the highly diverse group of college-enrolled individuals who do not graduate from a four-year college—a group commonly lumped together as "some college." We add to Heckman, Humphries, Veramendi, and Urzua (2014) by investigating this group in greater detail. Therefore, we distinguish between dropouts and graduates because we argue that the different tracks chosen and finished—two- or four-year college attendance (aspiration) and completion or non-completion (attainment)—may result in different effects on self-esteem.

We build on Kane and Rouse (1995), who investigate the labor market relevance of various educational outcomes by dividing the group of "some college" into four subgroups, depending on the type of college enrolled in and type of degree received: "only attended two-year college (no degree), only attended four-year college (no degree), attended both two- and four-year college (no degree), A.A. (highest degree)." However, in contrast to our approach, Kane and Rouse (1995) investigate college enrollment and degree attainment separately, i.e., their group "A.A." includes both two-year college graduates who never attended a four-year college and students transferring between two- and four-year colleges without obtaining a bachelor's degree. Therefore, they can not investigate the effect of failed educational aspirations. Nevertheless, Kane and Rouse (1995) show considerable labor market payoffs for only attending two- and four-year colleges, with an additional wage premium for completing an associate's degree. These differences in outcomes for the subgroups with some college experience but no bachelor's degree again show the importance of using a detailed classification when investigating the effect of education on various outcomes.

2.3 Data and Method

2.3.1 Classification of Education

Our framework for classifying final post-secondary schooling outcomes, our main explanatory variable, distinguishes among five college outcomes. Building on Kane and Rouse (1995), we focus on different educational paths: types of colleges attended in connection with final degree attainment.[2] To investigate the relationship between dropping out of college and self-esteem, we rely directly on reported degrees and connect them to the types of college. Using this approach, we determine five college outcomes.[3]

[2] For a less aggregated version of a similar approach fully relying on educational paths, see Agan (2014).

[3] Scholars commonly use completed years of schooling to determine two groups having at least some college experience (some college and college graduate). For example, see Wolpin (2005) for an explanation of this procedure for the NLSY79. However, in general this linear approach is not able to separate two-year college graduates and four-year college dropouts, as both can have the same number of years of schooling completed. Moreover, neither can the reverse approach—using actual degree obtainment to assign typical completed years of schooling (e.g., Park, 1996)—be used to classifying the un-ordered set of

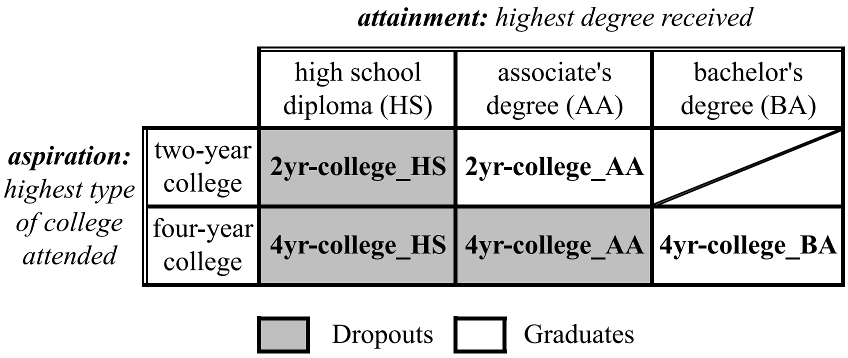

Figure 2.3.1: College Outcomes Using Attendance and Degrees

Figure 2.3.1 shows our five college outcomes. We classify each group by the highest type of college ever attended, i.e., two-year college (2yr-college) or four-year college (4yr-college), and by the highest degree ever received, i.e., high school diploma (HS), associate's degree (AA), bachelor's degree or higher (BA). By combining an individual's highest type of college ever attended and highest college degree ever received, we generate the following five groups: 2yr-college_HS, 2yr-college_AA, 4yr-college_HS, 4yr-college_AA, and 4yr-college_BA. By imposing a strict ranking for the types of college (*2yr-college* < *4yr-college*) and degrees ($HS < AA < BA$), and by only using each individual's highest type of college attended and highest degree received,[4] we generate five mutually exclusive groups. From this classification of educational outcomes we can derive a clear definition of dropouts: individuals for whom the highest degree received is lower than the degree that is granted by the highest type of college ever attended. Dropouts' attained education is lower than their aspired education, i.e., dropouts have failed aspirations. In sum, we

educational outcomes that allows for two- and four-year college enrollment and graduation. Later, we also estimate models with the conventional measure of education derived from completed years of schooling, to find similarities and differences in the results from using the conventional approach compared to the results from using our approach.

[4]For example, someone who went to several different two- and four-year colleges and who finally graduated with a BA would be labeled as "4yr-college_BA." If this same individual finally finished with only an AA, he or she would be labeled as "4yr-college_AA."

investigate three groups of dropouts: 2yr-college_HS, 4yr-college_HS, and 4yr-college_AA.

Our classification of education enables us to precisely define dropouts without directly modeling the dropout decision. Thus our approach requires no further assumptions while still clearly defining mutually exclusive educational groups. In the same manner, our approach is highly flexible and does not impose a strict time structure on the decision to drop out of college. More generally, our approach does not impose any restrictions on an individual's educational path: He or she can take any number of years to finish a specific degree,[5] can stop his or her college education, continue it at some later point in time, and finish with or without a degree,[6] or can switch between multiple institutions. Given this complexity of educational decisions, our classification of education provides us with a simplified yet clear-cut definition of dropouts.[7]

2.3.2 Measure of Self-Esteem

To measure self-esteem, we rely on the Rosenberg scale (RS), a ten item measure of global self-esteem (Rosenberg, 1965). Although the RS is relatively short, it has proven to be valid and reliable, making it an efficient tool for measuring global self-esteem (for an overview, see Goldsmith et al., 1996). Indeed, the RS is the most popular measure of global self-esteem among researchers in psychology and sociology (Baumeister et al., 2003; Blascovich & Tomaka, 1991; Rosenberg, Schooler, Schoenbach, & Rosenberg, 1995). Starting in the early 2000s, economists began to widely use both the RS and various

[5]For example, finishing a bachelor's degree within four years is far from being the U.S. norm (Jaeger & Page, 1996; Stratton & Wetzel, 2013).

[6]Arcidiacono, Aucejo, Maurel, and Ransom (2016) show that stopping-out of college (i.e. leaving college only temporarily) is a frequent occurrence, particularly at two-year colleges. Further, when investigating dropout behavior, Stratton, O'Toole, and Wetzel (2008) show the importance of using a classification of education that clearly does not include short-term "stopouts" in the group of long-term dropouts.

[7]However, our classification of education does not enable us to use regional or time variation for further analyses. Given the complexity of the dropout decision stretching over a long period of time and several colleges, we cannot attach each dropout to a certain year or a certain college. Including region or time information, we would be able to use random variation in dropout rates over region and time to estimate the causal effect of dropping out of college on self-esteem. However, as we cannot base our identification strategy on regional or time variation, we can only utilize the panel structure of our data to identify our effects.

adaptations of it (Bowles, Gintis, & Osborne, 2001; De Araujo & Lagos, 2013; Drago, 2011; Goldsmith et al., 1996, 1997; Heckman, Humphries, Urzua, & Veramendi, 2011; Heckman & Kautz, 2012; Heckman et al., 2006; Murnane et al., 2001; Persico, Postlewaite, & Silverman, 2004; Waddell, 2006).

The RS is a 10-item Likert scale (0=strongly disagree, 4=strongly agree) designed for measuring feelings of self-worth and self-acceptance (Blascovich & Tomaka, 1991). The statements are as follows: (1) I feel that I am a person of worth, at least on an equal basis with others; (2) I feel that I have a number of good qualities; (3) All in all, I am inclined to feel that I am a failure; (4) I am able to do things as well as most other people; (5) I feel I do not have much to be proud of; (6) I take a positive attitude toward myself; (7) On the whole, I am satisfied with myself; (8) I wish I could have more respect for myself; (9) I certainly feel useless at times; (10) At times I think I am no good at all.[8] Given various approaches for deriving the overall score, this paper follows the most common procedure—that of taking the sum of the items (Blascovich & Tomaka, 1991). This number ranges from 0 to 40,[9] i.e. the higher the score, the higher the self-esteem of the individual.

2.3.3 National Longitudinal Survey of Youth 1979

Our data comes from the NLSY79, a sample of 12,686 individuals born between 1957 and 1964 and first interviewed in 1979. The NLSY79 provides both the RS for 1980, 1987, and 2006, and detailed information on educational attainment. We construct a sample to empirically investigate the relationship between college enrollment and dropping out or graduating and self-esteem. This sample enables us to investigate the long-term effects of dropping out of college on self-esteem in 2006, when the individuals are in their 40s. Moreover, by giving individuals until 2006 to finish their education, we make our setting

[8]These items are developed in Rosenberg (1965). Items 3, 5, 8, 9, 10 are reverse-scored, for building the sum of the items.

[9]For our regression models we standardize all self-esteem scores to have a mean of 0 and a standard deviation of 1.

highly flexible, without strong restrictions on years in college or on leaving college and returning later. Therefore, we are able to observe individuals who have had sufficient time to finish all their post-high school education (for similar argumentation, see Kane & Rouse, 1995).

We impose several sample restrictions that enable us to classify an individual's college education in detail (for more details, see table 2.6.1). We restrict our sample to individuals still in the panel in 2006 and having reported a self-esteem score in 1980 and 2006. Moreover, we drop all individuals not enrolled in college at some point or for whom we do not have valid information on the type of college enrolled in or the highest degree obtained.[10] Finally, we remove observations with missing values on control variables. Our final sample consists of 2,836 observations.

Table 2.3.1 shows summary statistics for the main sample. The educational composition of the sample shows that the different college outcomes occur at very different rates with three larger groups (2yr-college_HS, 4yr-college_HS, and 4yr-college_BA) and two smaller groups (2yr-college_AA, 4yr-college_AA). In sum, 54.8 percent of the individuals are classified as dropouts (2yr-college_HS, 4yr-college_HS, and 4yr-college_AA). For four-year colleges only, 71 percent of the students attend such a college at one point, and 46 percent of these students never graduate from it.[11]

Table 2.3.1 shows that self-esteem increases on average from 1980 to 2006 for all

[10]For this reason, we also have to drop all individuals with the highest degree specified as "other" as long as they do not have indicated another highest degree in preceding or succeeding years. This restriction most likely affects two-year college students holding a diploma or certificate, as these categories are not part of the questionnaire. However, as we cannot drop the individuals who do not complete these programs, our group 2yr-college_HS may also include individuals who have only been enrolled in two-year college programs that do not grant an AA. In summary, we drop the graduates of these programs but have to keep the dropouts, and, therefore, we will most likely underestimate the dropout effect for two-year college-enrolled individuals.

[11]Although Light and Strayer (2000) report a dropout rate for four-year college students of 58 percent in their sample of the NLSY79, they only allow each individual to attend one college. Our approach does not restrict the number of colleges to which individuals might transfer, thereby resulting in a lower dropout rate. While our approach of classifying educational outcomes does not allow us to estimate a precise two-year college dropout rate, we can still provide a lower bound for this rate. By definition, the group 4yr-college_AA graduates from a two-year college at one point, and therefore the two-year college dropout rate in our sample must be at least 59 percent. We calculate this rate by dividing the number of individuals in the group 2yr-college_HS by the sum of the number of individuals in the groups 2yr-college_HS, 2yr-college_AA, and 4yr-college_AA.

college outcomes. In general, the descriptive results support our later findings on the negative effect of dropping out of college on self-esteem. In 1980 the ranking of the self-esteem score corresponds to a typical ranking of educational outcomes, which puts a strong emphasis on four-year college enrollment ($2yr - college\_HS < 2yr - college\_AA < 4yr - college\_HS < 4yr - college\_AA < 4yr - college\_BA$). However, in 2006 the group 2yr-college_AA ends up at a higher level of self-esteem than both types of four-year college dropouts (4yr-college_HS and 4yr-college_AA). This finding already indicates that merely entering but not completing a four-year college might have negative effects on the development of self-esteem.[12]

To describe our five educational groups in greater detail, we also report summary statistics for other background variables in Table 2.3.1. Investigating the AFQT scores—our most important background variable—we observe expected patterns. The group 2yr-college_HS has the lowest score, followed by three intermediate groups (2yr-college_AA, 4yr-college_HS, and 4yr-college_AA), and 4yr-college_BA as the highest scoring group. For the other background variables, we find that the group 4yr-college_BA differs from the other groups, which all have very similar statistics, in two ways. First, the group 4yr-college_BA has a lower percentage of black and hispanic individuals. Second, it differs in received family resources, i.e., the group has higher educated parents, higher family income, and less siblings. Interestingly, these differences exist only for graduates of four-year colleges but not for students who only enrolled and never graduated from a four-year college (4yr-college_HS and 4yr-college_AA). Therefore, differences in background variables appear to increase the likelihood of graduating from a four-year college but are unrelated to the likelihood of enrolling in one. In sum, these differences underscore the importance of controlling for all of these background variables when investigating the effects of educational outcomes.

[12]The correlation between the self-esteem score in 1980 and 2006 is 0.275. Thus, at least to some extent, self-esteem appears to develop differently across individuals.

Table 2.3.1: Summary Statistics

	2yr-college_HS	2yr-college_AA	4yr-college_HS	4yr-college_AA	4yr-college_BA	Total
N	618	196	698	229	1,095	2,836
%	21.8	6.9	24.6	8.1	38.6	100
Self-Esteem Score in 1980	22.23	22.99	23.03	23.38	23.94	23.23
	(3.85)	(3.84)	(4.02)	(4.15)	(3.88)	(3.98)
Self-Esteem Score in 2006	23.68	24.30	24.10	24.21	24.96	24.36
	(4.47)	(4.31)	(4.40)	(4.30)	(4.03)	(4.29)
AFQT Score	59.92	66.18	65.85	68.69	83.31	71.55
	(18.37)	(17.87)	(18.93)	(18.92)	(15.99)	(20.11)
Age (in 1979)	17.34	17.55	17.48	17.60	17.53	17.48
	(2.23)	(2.19)	(2.23)	(2.23)	(2.27)	(2.24)
Female	0.57	0.62	0.54	0.56	0.53	0.55
	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)	(0.50)
Black	0.30	0.22	0.36	0.29	0.19	0.27
	(0.46)	(0.41)	(0.48)	(0.46)	(0.39)	(0.44)
Hispanic	0.22	0.20	0.19	0.21	0.12	0.17
	(0.42)	(0.40)	(0.40)	(0.41)	(0.32)	(0.38)
Parents' Education	11.58	11.57	12.19	11.91	14.05	12.71
	(3.04)	(3.29)	(3.07)	(3.04)	(3.13)	(3.28)
Family Income (in 1979)	15.66	16.03	16.90	16.95	24.20	19.39
	(11.19)	(10.87)	(11.95)	(12.93)	(16.21)	(14.15)
Number of Siblings (in 1979)	3.64	3.89	3.59	3.75	2.87	3.35
	(2.36)	(2.77)	(2.56)	(2.49)	(2.05)	(2.37)
Urban Area Resident (in 1979)	0.79	0.78	0.81	0.79	0.81	0.80
	(0.41)	(0.41)	(0.39)	(0.40)	(0.40)	(0.40)

Notes: Reported are mean coefficients and standard deviations in parentheses. Columns with data on dropouts are shaded in gray. Region of residence in 1979 (Northeast, North Central, South or West) is not reported. The family income is divided by 1,000. The AFQT score is the sum of the scores on the Armed Forces Qualification Test conducted in 1980 (including tests on paragraph comprehension, word knowledge, mathematics knowledge, and arithmetic reasoning). The AFQT score was collected, as a part of the Armed Services Vocational Aptitude Battery, for all survey participants in the same year. The parents' education is equal to the father's or the mother's education, whichever is higher.

NLSY79, Authors' calculations.

2.3.4 Method

We treat our self-esteem measure as continuous and run OLS regressions.[13] To estimate the relationship between college outcomes and self-esteem, we use the following equation:

$$SE\_After_i = \alpha + \beta \cdot SE\_Before_i + \boldsymbol{\gamma_\eta} \cdot \mathbf{College\_outcome_i} + \boldsymbol{\delta_\lambda}\mathbf{X_i} + \varepsilon_i \tag{2.3.1}$$

We estimate the final level of self-esteem and control for the lagged measure. This approach is computationally equal to estimating the change while controlling for the lagged measure (for a discussion of different model specifications, see Allison, 1990). The variable SE_After_i represents the self-esteem score in 2006 (and 1987, for some additional analyses). SE_Before_i represents the lagged self-esteem score, measured in 1980. $\mathbf{X_i}$ is a vector of background variables measured in 1979 and 1980 (for details, see table 2.3.1). These baseline characteristics include parents' education; the Armed Forces Qualification Test (AFQT) score, which is generally viewed as a measure of skills that are important both in college and in the workplace (Light & Strayer, 2000);[14] and other time invariant reasons for dropping out of college. $\mathbf{College\_outcome_i}$ is a vector of η dummy variables, indicating the college outcome (see figure 2.3.1). $\boldsymbol{\gamma_\eta}$ is the vector of the main coefficients of interest.

We apply two strategies that enable us to limit concerns about selection issues. A selection bias arises because college outcomes are not randomly assigned: Instead, individuals decide to enroll in certain types of colleges, from which they can choose to graduate or to drop out. To reduce this bias, we apply the following two strategies. First, by focusing only on college-enrolled individuals, we limit the number of potential

[13] Prevoo and ter Weel (2015) apply a similar approach. As a robustness check, we also run ordered-probit regressions.

[14] Herrnstein and Murray (1994) view the AFQT score as an intelligence test, as it is a good measure of general cognitive ability. For further details, see also the discussion in Almlund et al. (2011). We follow Drago (2011) and standardize the score to have a mean of 0 and a standard deviation of 1 within each age group.

educational decisions and receive a more homogeneous sample, i.e. this restriction reduces the heterogeneity between the groups with different educational outcomes. Second, in addition to the lagged self-esteem measure, we include in our estimation other time invariant potential reasons for dropping out. An established reason for dropping out of college is family background (i.e., see Manski, 1992, for the effect of family income and parents' education, see Stinebrickner & Stinebrickner, 2008, for the effect of credit constraints on dropping out of college, and, see Light & Strayer, 2000, for the effect of mother's education on dropping out of college). Other reasons include general ability (AFQT score), race, and gender (Light & Strayer, 2000). To control for all these potential reasons for dropping out, we include them in $\mathbf{X_i}$.

To interpret our results as causal, we need to make relatively strong assumptions. To account for the influence of unobservable factors on the decision to drop out of college and on current self-esteem, we include lagged self-esteem and various control variables (including an ability measure, i.e., the AFQT score) in our model. To interpret our results as causal, we thus need to assume that dropouts and graduates would have the same increase in self-esteem over time, had the dropouts never dropped out. The main identifying assumption of our regression equation is independence of treatment status and self-esteem in 2006 conditional on self-esteem in 1980 and control variables. Specifically, all time invariant unobservable factors affecting both the self-esteem 2006 and education decisions are already captured in the 1980 self-esteem score and in the control variables. In sum, dropouts only differ systematically from graduates in 1980 self-esteem and in control variables but in no other dimension that might affect self-esteem in 2006. The degree to which this assumption is fulfilled determines the causality of our results.

2.4 Results

2.4.1 Main Results

Table 2.4.1 presents evidence on the negative effect of dropping out of college on self-esteem. Column (1) gives the results for the subsample of four-year college-enrolled students. The coefficient of the dummy variable 4yr-college_HS indicates that four-year college dropouts with only a high school degree score on average -.19*** standard deviations less on the self-esteem score in 2006 than four-year college graduates (4yr-college_BA), even when we control for differences in lagged self-esteem, various background variables, and cognitive ability. A similar effect applies to four-year dropouts who receive an associate's degree (4yr-college_AA).[15] In terms of self-esteem, obtaining an associate's degree is not a helpful strategy for preventing damage to self-esteem from dropping out of a four-year college. Put differently, having unfulfilled educational aspirations is detrimental for self-esteem.

In column (2), we show a similar pattern for the full sample, including also the students who enrolled exclusively in two-year colleges but never in a four-year college. The negative effect of two-year college dropouts (2yr-college_HS) is similar to the other dropout effects.[16] Furthermore, using one dummy variable indicating all kinds of dropouts (2yr-college_HS, 4yr-college_HS, and 4yr-college_AA), we find a coefficient for dropouts of -.16*** standard deviations. In summary, we find that all college dropouts have significant lower levels of self-esteem than four-year college graduates.

Therefore, our finding that two-year college graduates who never enrolled in a four-year college (2yr-college_AA) experience no such negative effect is striking.[17] Taken

[15]In column (1), we cannot reject the hypothesis that the coefficients of 4yr-college_HS and 4yr-college_AA are equal, with the probability value of the corresponding F-test being equal to .8384.

[16]Again, in column (2), we cannot reject the hypothesis that the coefficients of 2yr-college_HS, 4yr-college_HS, and 4yr-college_AA are equal, with the probability value of the corresponding F-test being equal to .8565. The hypothesis that all significant negative effects are equal cannot be rejected for all further tables, unless we report otherwise.

[17]This result is driven not by large standard errors but by a relatively smaller coefficient. Moreover, when we use the non-standardized self-esteem score in 2006 as the dependent variable, the OLS esti-

together, our results may indicate that merely obtaining any degree is not enough. The most important factor is not dropping out of the highest type of college ever enrolled in. More specifically, the differences between the group 2yr-college_AA and 4yr-college_AA show that self-esteem formation is not driven solely by the highest degree the student received but primarily by the highest type of college he or she enrolled in and whether he or she obtained a degree from that college. In this sense, our results are an indication of a dropout effect, not a graduation effect. Put differently, the development of self-esteem crucially depends on whether or not educational aspirations are fulfilled and not so much on the obtained educational degrees.

Table 2.4.1 also sheds light on the differences between our approach, which focuses on college attendance and degrees, and the more common approach of using completed years of schooling to classify college outcomes.[18] In column (3), we repeat our analysis using three college outcomes, derived by using not attendance and degrees but completed years of schooling: four-year college graduate, some college, and only started the first year of college without finishing it.[19] We find a strong negative effect for individuals with some college education (between 13 and 15 completed years of schooling). We also find a similar effect for the additional group of individuals who only started the 13th year of education. Our approach to classifying college outcomes generates qualitatively similar results. However, from comparing column (2) with column (3), we gain the additional insight that dropouts drive the effect for "some college," not two-year college graduates (2yr-college_AA). By deriving this result, we show that the fulfillment of educational aspirations is crucial for self-esteem development. In this regard, we add to the important previous findings in the literature by showing that unfulfilled educational aspirations are

mations of the models in columns (1) and (2) in Table 2.4.1 make no predictions outside of the logical range of values. Furthermore, the general result of a negative coefficient of the dummy variables indicating dropouts (2yr-college_HS, 4yr-college_HS, and 4yr-college_AA) holds when we use an ordered-probit specification to estimate the models of column (1) and (2) in Table 2.4.1.

[18]A similar argument applies when we use actual degree obtainment to assign typical completed years of schooling to the various outcomes.

[19]The last group cannot be identified when only using completed years of schooling to classify educational outcomes.

the driver of the effect for "some college."

Compared to the models based on years of schooling, our approach of using discrete college outcomes based on attendance and degrees has three additional advantages. First, we can include college dropouts who never finished their first year of college education. Second, we can test whether the effects for two-year college graduates and two- or four-year college dropouts are different (see table 2.4.1, column 2). Third, our approach enables us to make precise and detailed estimations of the effects for four-year college students (see table 2.4.1, column 1).

Although the magnitude of the effect for the dropouts appears not to be particularly large, the magnitude becomes of interest because self-esteem appears to be a trait-like characteristic that does not change much over the course of an individual's adult life. For example, one might interpret the coefficient of the self-esteem score in 1980 as the share of each standard deviation of the self-esteem score in 1980, which factors into the score in 2006. In columns (2) and (3), one additional standard deviation in 1980 results in about .26 standard deviations more in 2006, even when we control for a rich set of background variables that potentially could have already influenced self-esteem in 1980. This trait-like characteristic leaves only limited variation which could be explained with inputs between the two self-esteem measures. In light of this persistence over time, we view our estimated effects as relatively large.

Table 2.4.1: Explaining Self-Esteem Score in 2006 (OLS)

	4yr-college	Full	
	(1)	(2)	(3)
Self-Esteem Score in 1980	0.2404***	0.2628***	0.2617***
	(0.0223)	(0.0193)	(0.0194)
2yr-college_HS	-	-0.1772***	-
	(-)	(0.0548)	(-)
2yr-college_AA	-	-0.0726	-
	(-)	(0.0773)	(-)
4yr-college_HS	-0.1912***	-0.1512***	-
	(0.0527)	(0.0510)	(-)
4yr-college_AA	-0.1761**	-0.1398**	-
	(0.0723)	(0.0712)	(-)
Started Only First Year of College	-	-	-0.1926***
	(-)	(-)	(0.0590)
Some College	-	-	-0.1184***
	(-)	(-)	(0.0427)
Female	-0.1271***	-0.0987***	-0.1010***
	(0.0429)	(0.0364)	(0.0365)
Family Income (in 1979)	-0.0030*	-0.0019	-0.0019
	(0.0016)	(0.0014)	(0.0014)
AFQT Score	0.0126	0.0353	0.0358
	(0.0308)	(0.0256)	(0.0252)
Additional Background Variables	YES	YES	YES
Constant	2.8902**	2.3826**	2.3649**
	(1.3843)	(1.1935)	(1.1932)
R-squared	0.088	0.095	0.095
Observations	2,022	2,836	2,836

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Rows with effects for dropouts are shaded in gray. The base group in all columns is 4yr-college_BA. Additional background variables are age (in 1979), age (in 1979) squared, dummy variables for race, dummy variables for region of residence in 1979, number of siblings, urban area resident (in 1979).

NLSY79, Authors' calculations.

2.4.2 Additional Results and Robustness Checks

This subsection presents additional results showing the robustness of our findings. First, we show the robustness of our results with regard to selection issues in education. Second, we show when the effects emerge and how they potentially fade out over time. Third, we show that our results are not purely driven by different labor market outcomes or other confounding factors in which dropouts and graduates may differ.

Selection Issues

To test the robustness of our results with regard to selection issues in education, we repeat our analysis for a subsample of college students whose lagged self-esteem score was measured while they were still enrolled in high school. Doing so ensures that the self-esteem measure in 1980 is not affected by prior or current schooling decisions. To ensure that the lagged self-esteem is not already affected by educational outcomes, we restrict our sample to individuals who are still enrolled in high school in 1980. Our identification relies on the assumption that the lagged self-esteem measure in 1980 is not affected by educational outcomes. In 1980, some individuals might still be enrolled in high school while others have graduated, and started college, or even finished college. Restricting our sample to individuals who were still enrolled in high school in 1980 reduces the effect of current or past education on self-esteem. Therefore, we construct a subsample in which all individuals are still enrolled in high school in 1980. This restriction decreases our sample size from 2,836 to 1,227 observations.

In general, our results are similar to our main findings for both only four-year college students and for the full sample (table 2.6.2). We find a persistent gap in self-esteem between college dropouts and four-year college graduates. Again, we find no negative effect for the group 2yr-college_AA. In magnitude, our results are larger than those for our primary investigation in Table 2.4.1. Although, the differences in magnitude might be driven by the different sample composition, our primary results are not qualitatively affected by the sample composition.

Time Structure

Given the long time period between our two self-esteem measures (26 years), we investigate the time structure in more detail. We provide evidence on the early emergence and fading out of our effects by explaining the self-esteem score measured in 1987; seven years after our initial self-esteem measure. In 1987, many individuals in our sample were still enrolled in college. Therefore, to ensure that ongoing education does not affect the self-esteem score, we include a subsample analysis restricted to individuals who finished their education by 1987 and did not enroll in college until at least 2006. This restriction decreases our sample size from 2,836 to 1,501 observations.

Using the self-esteem score in 1987, we find very similar results for four-year college students: dropping out of a four-year college reduces self-esteem, regardless of whether the dropouts received an associate's degree or not (table 2.6.3, columns 1 and 3). However, the magnitude of the effects is bigger than in Table 2.4.1. In particular, the results for the subsample analysis in Table 2.6.3 column (3) reveal much higher negative effects of dropping out of college on self-esteem when self-esteem is measured close to the time that college education ends. However, while the negative effect of dropping out of college might fade out over time, we know that it remains still negative and significant at least until 2006 (see table 2.4.1).

In the sample that includes also two-year college students, the patterns become less clear. While the group 2yr-college_HS still experiences a negative effect of a similar magnitude, the group 2-yr-college_AA now also experiences a significant negative effect (in contrast to our results in table 2.4.1). While this result appears puzzling, it could represent yet another facet of the broader idea that the effects of educational experience on self-esteem fade over time. When we look at the summary statistics in Table 2.3.1, we see that self-esteem scores for all groups appear to converge between 1980 and 2006, that is, the spread between the respective group with the highest and the lowest self-esteem score decreases from 1 to .68 points on the self-esteem scale. In some sense this reduced range could also explain the observed differences between our results in Tables 2.4.1 and 2.6.3: Differences between the groups appear to decrease over time (regression to the mean). Not only the negative effects for the dropouts but also the positive effects for the base group, 4yr-college_BA, decrease. This development may explain why we observe a significant negative effect for two-year college graduates when we investigate the self-esteem score in 1987 but not when we investigate it in 2006.

Mechanisms and Confounders

Dropouts and graduates might also differ in various other dimensions which might confound our results. To understand how self-esteem emerges over time and through what channels dropping out has a negative effect on self-esteem, we investigate whether our effects disappear after we control for (intermediate) outcomes of education. We therefore estimate a more direct model, compared to the reduced form model that we report in Table 2.4.1. In this direct model we include several potential outcomes of college education or other confounding factors measured in 2006: income, labor force status, number of children, and martial status. While such a model might be over-controlled, it still can help us to understand the potential mechanisms driving our results.

When we include these additional variables measured in 2006, the effect of dropping out on self-esteem remains significant for four-year college students (table 2.6.4, column 1). Moreover, the effect remain similar in magnitude to the results in Table 2.4.1. In Table 2.6.4 column (2) we run the same regression for the full sample. For the groups 2yr-college_HS and 4yr-college_HS the negative effect still exists. However, the group 4yr-college_AA no longer experiences the negative effect on self-esteem. If the group 4yr-college_AA had the same labor market outcomes as the group 4yr-college_BA (hypothetical), they would not experience any negative effect on self-esteem. This finding appears to suggest that wage expectations and reference group effects might be important for the group 4yr-college_AA. In sum, had they been able to offset the effect of a bachelor's degree on the intermediate outcomes we control for (e.g., when they would earn the same wages as graduates), they would not experience a reduction in self-esteem. In contrast, for the other groups, we find lower levels of self-esteem even when holding outcome variables fixed. Therefore, the effect of dropping out of college on self-esteem is not entirely driven by the fact that dropouts and graduates differ in outcome variables measured in 2006. In other words, an additional effect of dropping out on self-esteem exists, beyond the potential effect of different labor and non-labor market outcomes for dropouts and graduates.

2.4.3 Economic Importance of Results

To attach an economic value to our results, we perform a back-of-the-envelope calculation of the effect of self-esteem on wages.[20] Table 2.6.5 shows the results for typical regressions of self-esteem on wages. The returns to a one standard deviation increase in self-esteem 2006 range between 5.3 and 11.0 percent and are thus in line with the results of an IV estimation by Drago (2011). Column (4) shows that self-esteem is valued in the labor market, even when controlling for a large set of controls, including education. Therefore, to derive a rough estimate of the economic value of our results, we use the returns of 6.9 percent in Table 2.6.5 column (3) because this model represents as closely as possible our main model for the development of self-esteem.

The estimates in Table 2.6.5 show that self-esteem has a significant effect on wages. We combine these returns (6.9 percent) with our main effects on self-esteem of being a dropout, which range between -.14 and -.18 standard deviations (table 2.4.1, column 2), with the effect for the entire group of dropouts being -.16 standard deviations. Using the returns of 6.9 percent and the effect size of -.16 standard deviations, our effect of dropping

[20]For this calculation we restrict the sample to individuals with valid data on wages and tenure in 2006.

out is equal to -1.1 percent, that is, the wages of dropouts and graduates differ by about 1.1 percent due to their different levels of self-esteem. However, these differences are part of observed differences in traditional returns to education. In this sense we provide an insight into exactly what constitutes returns to education. As self-esteem is valued in the labor market, and as dropping out of college results in significantly lower levels of self-esteem, the observed wage differences between dropouts and graduates can be explained—to some extent—by differences in self-esteem.

2.5 Conclusion

Using the NLSY79, our paper investigates the long-term relationship between unfulfilled educational aspirations and self-esteem. Therefore, we develop a classification to separate college dropouts with unfulfilled aspirations and graduates with fulfilled aspirations. Compared to graduating, attending a four-year college and not finishing it results in a significant lower level of self-esteem. Similarly, attending a four-year college but ending up only with an associate's degree results in lower self-esteem, just as attending a two-year college but not finishing it. However, those who both exclusively attend and finish a two-year college with an associate's degree do not experience any such effect on self-esteem. Thus our results suggest that in terms of developing self-esteem, avoiding dropping out of any education is most critical as ending up with unfulfilled educational aspirations has long-term negative effects. Therefore, aiming for an education that has a high likelihood of success could be a good strategy for avoiding damage to a student's self-esteem.

The results in this paper are limited in at least two ways, both of which provide potentials for future research. First, by including the possibility of dropping out of college several times in the construction of our college outcomes, we are not able to use information on individual colleges that includes information on either college quality, or differences in dropout rates over regions, or over time, or over both. Dropping out of a low-quality college might affect an individual's self-esteem less than dropping out of a high-quality one, although the opposite might also hold true. However, as these effects might differ over the distribution of college quality, further analysis in this direction could be beneficial.

The second limitation results from data constraints, based on the timing of the survey interviews. The periods between our self-esteem measures might be sufficiently long that several events, unrelated to schooling, might also systematically alter self-esteem. In our robustness checks we address this issue and also investigate the short-run effects, which appear to be much larger than the long-term effects. However, having data on self-esteem at the college enrollment date, at a date closely following either the dropping out or the graduating, and long-run follow-ups, researchers would be able to address this issue even more closely. Self-esteem assessments more closely following the end of college education might particularly provide a good opportunity for estimating the—potentially very large—immediate effects more precisely. Doing so, however, calls for future cohorts and survey waves.

Our results, which we derive by using college outcomes based on attendance and degree attainment, imply that associate's degrees are an efficient means of developing self-esteem. A student's completion of a two-year college, when he or she has never been enrolled in a four-year college, yields an increase in self-esteem equal to the increase for students completing a four-year college. Dropouts from two- and four-year colleges miss out on this effect. Therefore, our results for the development of self-esteem echo the results of Kane and Rouse (1995) for the financial returns to different types of colleges: finishing a two-year college outperforms merely attending a four-year college. Indeed, our results even go one step further by showing that enrolling in a four-year college and failing to finish it is detrimental for self-esteem, regardless of whether obtaining an associate's degree or not.

Thus students need to form realistic educational aspirations. Therefore, career counseling should help to form these aspirations by providing information on the demands of different types of colleges. Moreover, students might well be advised to consider the strategy of enrolling first in institutions from which they can reasonably expect to graduate. For example, students, who are not sure whether they can succeed at a four-year college, could be advised to first enroll in a two-year college, from which they can expect to graduate. As a result of the two-year college experience, they will be able to make better-informed decisions about four-year college enrollment. In this regard, educators may facilitate such a step-wise process by increasing the permeability between different educational institutions. Thus our findings have implications for both educators and students alike.

2.6 Appendix

2.6.1 Analytic Sample

Table 2.6.1: Sample Construction

	N
Intial sample (NLSY79)	12,686
In panel in 2006	7,654
Self-Esteem not missing	7,074
College enrolled with valid information	3,744
Controls 1979 not missing	2,942
Controls 2006 not missing	2,836
Analytic sample	**2,836**

Notes: Column 1 presents the restrictions applied to create the sample. Column 2 shows the number of observations left after each sample restriction.

NLSY79, Authors' calculations.

2.6.2 Robustness Checks: Selection, Timing, and Confounders

Table 2.6.2: Explaining Self-Esteem Score in 2006 for Subsample Being Enrolled in High School in 1980 (OLS)

	4yr-college	Full
	(1)	(2)
Self-Esteem Score in 1980	0.2335***	0.2531***
	(0.0352)	(0.0298)
2yr-college_HS	-	-0.2686***
	(-)	(0.0787)
2yr-college_AA	-	-0.0845
	(-)	(0.1265)
4yr-college_HS	-0.2059**	-0.1797**
	(0.0827)	(0.0795)
4yr-college_AA	-0.3253***	-0.2975**
	(0.1183)	(0.1155)
Female	-0.1175*	-0.0918*
	(0.0663)	(0.0551)
Family Income (in 1979)	-0.0039	-0.0037
	(0.0029)	(0.0026)
AFQT Score	0.0010	0.0116
	(0.0471)	(0.0392)
Additional Background Variables	YES	YES
Constant	-5.0062	0.9905
	(6.6917)	(5.3271)
R-squared	0.095	0.095
Observations	849	1,227

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Rows with effects for dropouts are shaded in gray. The base group in all columns is 4yr-college_BA. Additional background variables are age (in 1979), age (in 1979) squared, dummy variables for race, dummy variables for region of residence in 1979, number of siblings, urban area resident (in 1979).

NLSY79, Authors' calculations.

Table 2.6.3: Explaining Self-Esteem Score in 1987 (OLS)

	All		Finished Education in 1987	
	4yr-College	Full	4yr-College	Full
	(1)	(2)	(3)	(4)
Self-Esteem Score in 1980	0.3822***	0.3737***	0.3825***	0.3655***
	(0.0215)	(0.0187)	(0.0304)	(0.0255)
2yr-college_HS	-	-0.1335**	-	-0.1480**
	(-)	(0.0525)	(-)	(0.0700)
2yr-college_AA	-	-0.2567***	-	-0.3750***
	(-)	(0.0739)	(-)	(0.1042)
4yr-college_HS	-0.2662***	-0.2622***	-0.3242***	-0.3136***
	(0.0482)	(0.0468)	(0.0674)	(0.0647)
4yr-college_AA	-0.1729**	-0.1680**	-0.2742**	-0.2650**
	(0.0711)	(0.0700)	(0.1204)	(0.1185)
Female	-0.1111***	-0.1125***	-0.0642	-0.0717
	(0.0410)	(0.0350)	(0.0572)	(0.0473)
Family Income (in 1979)	-0.0005	-0.0011	-0.0005	-0.0017
	(0.0015)	(0.0014)	(0.0020)	(0.0018)
AFQT Score	0.1361***	0.1386***	0.1428***	0.1418***
	(0.0290)	(0.0243)	(0.0395)	(0.0325)
Additional Background Variables	YES	YES	YES	YES
Constant	2.4894*	1.5957	4.7514**	1.8913
	(1.2946)	(1.1287)	(1.8425)	(1.5589)
R-squared	0.212	0.202	0.232	0.205
Observations	1,956	2,752	978	1,501

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Rows with effects for dropouts are shaded in gray. The base group in all columns is 4yr-college_BA. Additional background variables are age (in 1979), age (in 1979) squared, dummy variables for race, dummy variables for region of residence in 1979, number of siblings, urban area resident (in 1979).

NLSY79, Authors' calculations.

Table 2.6.4: EXPLAINING SELF-ESTEEM SCORE IN 2006 WITH ADDITIONAL CONTROL VARIABLES MEASURED IN 2006 (OLS)

	4yr-College	Full
	(1)	(2)
Self-Esteem Score in 1980	0.2304***	0.2533***
	(0.0221)	(0.0191)
2yr-college_HS	-	-0.1289**
	(-)	(0.0547)
2yr-college_AA	-	-0.0258
	(-)	(0.0768)
4yr-college_HS	-0.1478***	-0.1013**
	(0.0522)	(0.0505)
4yr-college_AA	-0.1495**	-0.1090
	(0.0724)	(0.0712)
Female	-0.0753*	-0.0413
	(0.0446)	(0.0378)
Family Income (in 1979)	-0.0033**	-0.0024*
	(0.0016)	(0.0014)
AFQT Score	-0.0020	0.0178
	(0.0309)	(0.0257)
Additional Background Variables	YES	YES
Control Variables and Labor Market Outcomes (2006)	YES	YES
Constant	2.4300*	1.9697*
	(1.3773)	(1.1829)
R-squared	0.110	0.118
Observations	2,022	2,836

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Rows with effects for dropouts are shaded in gray. The base group in all columns is 4yr-college_BA. Additional background variables are: age (in 1979), age (in 1979) squared, dummy variables for race, dummy variables for region of residence in 1979, number of siblings, urban area resident (in 1979). Control variables and labor market outcomes (2006) are income (in 2006), dummy variables for labor force status (in 2006), number of children (in 2006), dummy variables for martial status (in 2006).

NLSY79, Authors' calculations.

2.6.3 Self-Esteem and Wages

Table 2.6.5: Explaining Wages with Self-Esteem (OLS)

	(1)	(2)	(3)	(4)
Self-Esteem Score in 2006	0.1100***	0.0739***	0.0685***	0.0532***
	(0.0143)	(0.0130)	(0.0138)	(0.0130)
Self-Esteem Score in 1980	-	-	0.0563***	0.0432***
	(-)	(-)	(0.0149)	(0.0139)
Education, Tenure, Age, Age Squared	-	YES	-	YES
Background Variables	-	-	YES	YES
R-squared	0.027	0.202	0.199	0.294
Observations	2,331	2,331	2,331	2,331

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Dependent variable is log of hourly wages in 2006. Background variables are dummy variable for gender, family income (in 1979), AFQT score, age (in 1979), age (in 1979) squared, dummy variables for race, dummy variables for region of residence in 1979, number of siblings, urban area resident (in 1979). Education is measured in years of finished education.

NLSY79, Authors' calculations.

Chapter 3

Development of Non-Cognitive Skills during Apprenticeship Training

A short and revised version of this chapter is published in *Economics Letters* (163, 40-45) as "Development of Non-Cognitive Skills in Adolescence" by Hoeschler, Balestra, and Backes-Gellner (2018). Also, part of this chapter is a revised version of early parts of the working paper "Development of Non-Cognitive Skills in Adolescence" by Hoeschler, Balestra, and Backes-Gellner (2017).

3.1 Introduction

Economists recognize the importance of non-cognitive skills for individuals' economic behavior and their labor market outcomes (Heckman et al., 2006), a recognition that has led researchers to investigate the formation and stability of non-cognitive skills. While most of the literature suggests that non-cognitive skills are stable during adulthood (Cobb-Clark & Schurer, 2012), hardly any evidence exists on the formation and malleability of these skills during adolescence. This lack of research is surprising, because many personality traits appear responsive to policy interventions and educational investments (Almlund et al., 2011). However, both measurement issues and the lack of high-quality data have prevented researchers from providing definitive answers. Our paper intends to fill this research gap.

To study the formation and malleability of non-cognitive skills, we measure adolescents' non-cognitive skills with well-established multiple-question inventories at two points in time. We focus on two types of non-cognitive skills: Grit (i.e., the passion and perseverance towards long-term goals) and the Big Five personality traits (i.e., conscientiousness, extraversion, agreeableness, openness, and emotional stability). While studying the Big Five has a long tradition in economics and psychology, research on Grit started only in 2007 with the establishment of a reliable scale (Duckworth, Peterson, Matthews, & Kelly, 2007). Grit has attracted researchers' interest, because studies have shown it to be highly predictive of labor market success and other life outcomes (Eskreis-Winkler, Shulman, Beal, & Duckworth, 2014). As early as 2009—closely following the introduction of the Grit scale—we started a longitudinal data set that allows us to study the development of Grit among adolescents over a six-year period. In our analysis, we focus on the development of Grit and compare it to that of the Big Five.

This chapter contributes to the literature on non-cognitive skills by answering the following four questions: (1) Do Grit and the Big Five change during adolescence? (2) Are changes in these non-cognitive skills heterogeneous across individuals? (3) Can traditional education production function approaches explain changes in non-cognitive skills? (4) How much of the change is reliable under reasonable assumptions on measurement error?

3.2 Non-Cognitive Skill Development and Workplace-Based Education

Psychologists who have intensively studied the development of personality over the life span suggest that adolescence and young adulthood are critical phases in that development (Costa & McCrae, 1994; Donnellan, Hill, & Roberts, 2015; Roberts, Walton, & Viechtbauer, 2015). During these phases, individuals are usually enrolled in education and training programs. Given that psychologists also show that work-related experience is critical for the development of personality at this stage (Roberts, Caspi, & Moffitt, 2003; Roberts, Wood, & Caspi, 2010), we investigate the development of non-cognitive skills during apprenticeship training, which is a vocational education and training program combining education and work-related experience.

Vocational education and training, the main path of education in many parts of western Europe, is becoming increasingly important in both the UK and the U.S. We use a sample of adolescents enrolled in apprenticeship training programs in Switzerland, a country where 70 percent of every cohort chooses vocational education and training (Hoffman & Schwartz, 2015). The dual apprenticeship training program, which starts around age 16, lasts three to four years and combines one to two days per week of classroom learning in a vocational school with workplace-based training in a host company (Wolter & Ryan, 2011).

Moreover, apprenticeship training appears promising for the development of non-cognitive skills. Apprenticeship training has at least three characteristics that may foster the effective development of non-cognitive skills in adolescence: it is well-structured, it provides close mentoring, and it has labor-market proximity (Heckman & Mosso, 2014; Kautz, Heckman, Diris, ter Weel, & Borghans, 2014). Moreover, some first empirical indications point towards the role of vocational education and training for the development of non-cognitive skills. A study for Switzerland shows that workplace-based vocational education and training, i.e, apprenticeship training, affects the development of emotion-centered coping, a specific non-cognitive skill (Bolli & Hof, 2014). Differentiating between workplace- and school-based learning environments, another Swiss study shows that workplace-based education has a comparable advantage in developing non-cognitive skills (Bolli & Renold, 2017). These research findings lead us to investigate the development of Grit and the Big Five during apprenticeship training.

3.3 Data

3.3.1 Leading House Apprenticeship Panel

To study the development of non-cognitive skills during adolescence, we use a unique panel data set—the Leading House Apprenticeship Panel—which provides us with repeated measures of Grit and the Big Five between the ages of 15 and 22. The Leading House Apprenticeship Panel is a self-collected data set on a homogeneous cohort of 265 apprentices in three occupations (one commercial and two technical) in the canton of Zurich. The apprentices in the Leading House Apprenticeship Panel were first interviewed in their first week of apprenticeship in 2009, and then followed annually until after they finished their training and entered the labor market in 2014/2015. Both the initial survey in 2009 and the final survey in 2014/2015 include the same measures for Grit and the Big Five, thereby enabling us to investigate changes in these non-cognitive skills over time.

The panel data set is fully funded by the Swiss State Secretariat for Education, Research and Innovation (SERI) through its Leading House on the Economics of Education, Firm Behavior and Training Policies, and earlier waves of the data set have already been used in three projects. The data set was initiated in 2009 by our colleague Donata Bessey, who in one project of her dissertation, used the first wave of the panel to investigate the effect of non-cognitive skills on the sureness to graduate from apprenticeship training

(Bessey, 2010). Subsequent waves were collected by our colleague Yvonne Oswald, who used them in her dissertation to investigate the effects of risk attitude and time preferences during apprenticeship training (Oswald, 2013). Her first study investigates the connection between experimentally elicited time preferences and real-world dropout decisions, as well as later labor market outcomes (Backes-Gellner, Herz, Kosfeld, & Oswald, 2018; Oswald, 2013). Her second study investigates the effect of randomly assigned bonus schemes for school performance on school grades, focusing in particular on the interaction effect of bonuses and time preferences (Oswald & Backes-Gellner, 2014). In 2014/2015, we conducted the final survey, a full two years after the apprentices both finished their training and entered the labor market.

3.3.2 Outcome Measures

We derive the measures of non-cognitive skills from standard multiple-question inventories. For Grit, we use the 8-item Grit scale, a highly efficient full-sentence questionnaire developed in Duckworth and Quinn (2009). Our data set is the first to provide repeated Grit measures over six years, i.e., a period much longer than the one year reported in the literature (Duckworth & Quinn, 2009). This data set, thus is the first to allow the study of the development of Grit for such a long time in general and for adolescents in particular.[1]

For the Big Five, we use the established 3-items-per-trait scale, based on the original Big Five Inventory (BFI) scale and further developed in Gerlitz and Schupp (2005). To assess the Big Five, the BFI scale uses short-sentence questions, which John, Naumann, and Soto (2008) argue are less ambiguous and provide more consistent answers than the frequently used sets of adjectives (e.g., in Cobb-Clark & Schurer, 2012, and, Elkins et al., 2016).

[1]Even though Grit characterizes the passion for long-term goals, we do not view is a measure of time preferences. Indeed, the correlation between the Grit measure and experimentally elicited time preferences in our data is only .1.

To account for the different number of items and scales, each outcome measure is adjusted so that it ranges between 0 and 1. Therefore, to make the measures more comparable, we apply a linear transformation and divide each measure by its respective maximum. Grit is measured on an eight-item Likert scale with each item taking values between zero and four (with a maximum of 32), while the Big Five are measured on a three-item Likert scale, with each item taking values between zero and six (with a maximum of 18).[2] However, for most of the Big Five traits, the desired direction is not clear, that's whether more of a trait is always better. The one Big Five trait that clearly is "negative" is neuroticism and clearly less of that trait than more is desirable. Therefore, we follow the common approach of recoding neuroticism as its reverse, labeling it "emotional stability" (e.g., Cobb-Clark & Schurer, 2012), an approach that simply reverses the signs of our coefficients while in no way affecting our results. For further details, see the full questionnaire in the additional material at the end of this dissertation.

Table 3.3.1, column (1), provides summary statistics for our outcome measures. The initial (2009) and final (2015) measures range between .49 and .76. Moreover, all of these measures have fairly similar standard deviations. However, when comparing changes, we still divide these changes by the standard deviations of the baseline score. This approach enables us to express changes in standard deviation units, thus making them fully comparable. In addition, for each non-cognitive skill the standard deviation does not differ substantially between the two respective measures over time. This finding shows that our group of individuals is not becoming more homogeneous over time in terms of their non-cognitive skills.

[2]For agreeableness, we can only use two items due to data restrictions. Therefore, the maximum for agreeableness is 12.

Table 3.3.1: Summary Statistics of Non-Cognitive Skills

	Descriptive Statistics (1)				Estimated Change (2)	Correlation with Score 2009 (3)
	mean	sd	min	max	in 2009 sd	coefficient
Grit						
2009	0.599	0.133	0.188	0.906		
2015	0.666	0.136	0.344	1.000		0.2501**
Change	0.067	0.165	-0.344	0.812	0.5022***	-0.6039***
Conscientiousness						
2009	0.651	0.175	0.111	1.000		
2015	0.742	0.153	0.278	1.000		0.3498***
Change	0.090	0.188	-0.389	0.556	0.5163***	-0.6479***
Extraversion						
2009	0.692	0.214	0.000	1.000		
2015	0.700	0.214	0.111	1.000		0.5879***
Change	0.008	0.198	-0.500	0.556	0.0353	-0.4495***
Agreeableness						
2009	0.699	0.194	0.000	1.000		
2015	0.764	0.176	0.250	1.000		0.4610***
Change	0.065	0.193	-0.583	0.583	0.3368***	-0.5871***
Openness						
2009	0.622	0.181	0.167	1.000		
2015	0.622	0.172	0.222	0.944		0.4216***
Change	0.000	0.190	-0.611	0.611	0.0020	-0.5699***
Emotional Stability						
2009	0.492	0.198	0.056	1.000		
2015	0.571	0.192	0.111	0.944		0.4186***
Change	0.079	0.211	-0.500	0.556	0.3991***	-0.5602***

Notes: N=153. *** $p < 0.01$. The Grit measure is the sum of eight Likert scale items (0-4) divided by 32. The agreeableness measure is the sum of two Likert scale items (0-6) divided by 12. Each of the other Big Five measures is the sum of three Likert scale items (0-6) divided by 18. Emotional stability is calculated as the reverse of neuroticism (see page 43).

Leading House Apprenticeship Panel, Authors' calculations.

3.3.3 Attrition Analysis

Attrition is a major threat to all panel data sets. As our sample is not representative, attrition is no threat to the external validity of our results, that is, attrition does not affect the generalizability of our results. However, given our research question, attrition might be a threat to the internal validity of our results. For example, if only the individuals with a positive development in non-cognitive skills stay in the sample, we would obtain biased results. However, while we cannot test this issue directly, we can investigate whether attrition is related to our baseline scores.

We collected baseline measures at age 15-16 and final outcomes at age 21-22. The initial sample in 2009 consisted of 265 individuals, of whom 255 provided measures of non-cognitive skills. In the final wave, six years later, 159 individuals responded to our survey efforts (via e-mails, letters, phone calls, and social media), 153 of whom provided measures of non-cognitive skills. In sum, the overall attrition over the six-year period is 40 percent.

Table 3.8.7 shows that the attrition is unrelated to baseline non-cognitive skills. The table reports a standard attrition analysis in which we use the baseline non-cognitive skills to explain the probability of being in the sample in the final wave. Running various model specifications (OLS and Probit, both with and without background variables as additional controls), we find no statistically significant relationship between the probability of being in the sample and baseline non-cognitive skills. In addition, when looking at the raw correlations between each individual non-cognitive skill and the probability of being in the sample, we likewise find no significant relationships. Therefore, we conclude that the attrition is unrelated to baseline non-cognitive skills.

As our research design allows us to investigate those individuals who leave the survey in greater detail, we provide further information on them. Individuals leave our sample for two reasons: dropping out of the apprenticeship program (*dropouts*), i.e., not continuing the program that they began in 2009, or simply leaving the survey (*non-responders*).[3] While both groups leave the survey, they do so for very different reasons and therefore might have substantially different baseline non-cognitive skills. Table 3.8.4 shows the mean differences in baseline non-cognitive skills for these two groups.[4] We find that dropouts have significant lower levels of Grit and conscientiousness than non-responders. This result is in line with the consensus in the literature that these non-cognitive skills are important for educational success (e.g., Duckworth et al., 2007). Moreover, dropouts score higher on openness than non-responders, possibly providing another reason for their dropout decision. In sum, while finding significant differences in non-cognitive skills between dropouts and non-responders, we still show that leaving the survey is unrelated to non-cognitive skills.

3.4 Average Within-Person Changes

To examine individuals' changes in Grit and the Big Five over time, we start by investigating average within-person changes. Table 3.3.1, column (1), provides means and standard deviations for the baseline score, the final score, and the within-person change over time. We estimate a significant average increase for Grit of about .5 standard deviation units of the baseline score (table 3.3.1, column 2). For the Big Five, we estimate similar significant increases for conscientiousness and—to a lesser extent—for agreeableness and emotional stability.

To investigate these average changes in greater detail, we move towards a distributional analysis. For this distributional analysis, we again restrict our sample to individuals

[3]In total 102 individuals leave the data set: 82 non-responders and 20 dropouts. To identify the dropouts, we rely on information from a cantonal (state) office, the "Mittelschul- und Berufsbildungsamt" (MBA) in Zurich (for further details, see Oswald, 2013). After dropping out, individuals can either start a new apprenticeship program or entirely leave the apprenticeship system. In Switzerland, most dropouts who stay in the apprenticeship system start a new apprenticeship training in another occupation (BFS, 2017).

[4]For the full sample of the Leading House Apprenticeship Panel, Oswald (2013) provides an in-depth analysis of the dropout decision, with a focus on the role of time preferences. She shows that more patient students are less likely to drop out.

with valid measures of non-cognitive skills at both points in time. Figure 3.4.1 depicts the probability density functions at the two points (2009 and 2015). For Grit, Figure 3.4.1 shows that the mean-level increase is the result of a parallel shift of the distribution to the right. These developments could be the result of either of two mechanisms, both of which we will investigate: first, a homogeneous increase (i.e., each individual develops in the same positive way) and, second, heterogeneous changes over time (i.e., individuals experience both positive and negative changes, with the positive outweighing the negative on average). For the Big Five, we observe similar developments for conscientiousness, agreeableness, and emotional stability.

By investigating the explanatory power of the first potential mechanism, a homogeneous increase, we find that the mean-level changes are not the same for all individuals. While the mean-level change in Grit is sizable, the two scores over time appear only loosely correlated at .3 (table 3.3.1, column 3).[5] For the three Big Five traits with significant mean-level changes, we find very similar results (table 3.3.1, column 3). Moreover, while openness shows no mean-level change, it appears unstable at the individual level (table 3.3.1, column 3), indicating that while individuals experience positive and negative changes, these changes on average cancel one another out. In sum, Grit and some of the Big Five, primarily conscientiousness and—to a lesser extent—agreeableness and emotional stability change during adolescence, with significant average within-person increases. Further, given the relatively low correlations between the measures over time, we can rule out the possibility that all individuals experience a homogeneous increase over time. These findings leave us with the second explanation: that individuals experience heterogeneous changes over time. We investigate this potential explanation in the next section.

[5] Rank correlations over time are close in magnitude to the correlation coefficients in Table 3.3.1, column (3).

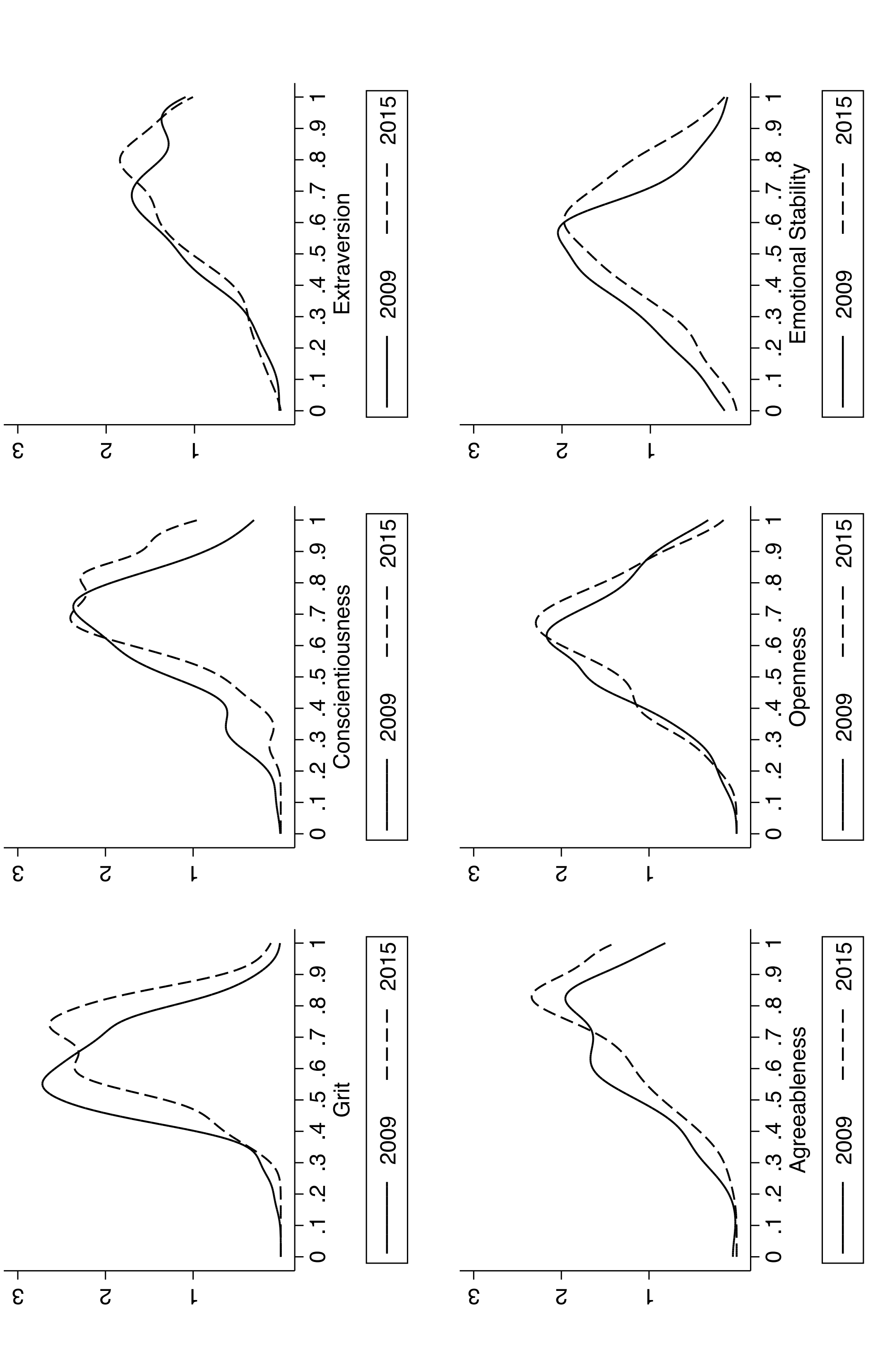

Figure 3.4.1: Probability Density Functions of Outcome Variables for Balanced Panel of Individuals
Notes: Leading House Apprenticeship Panel, Authors' calculations.

3.5 Heterogeneous Within-Person Changes

To investigate the heterogeneity in within-person changes, we depict the density functions of the within-person changes over time in Figure 3.5.1. For Grit, we clearly find that the development is heterogeneous during adolescence. Although the density mass of the changes is positive, some individuals show negative growth in Grit. This finding is also supported by Table 3.8.1, which represents a numerical analysis of Figure 3.5.1 by providing the percentage of individuals who—in absolute values—change (1) below one standard deviation, (2) between one and two standard deviations, and (3) more than two standard deviations. For Grit, 43 percent change at least by one standard deviation, and 12 percent change even by two standard deviations or more. Therefore, we show that Grit develops very heterogeneously, with both substantial positive and negative changes.

The general pattern of heterogeneity in changes also holds for the Big Five (figure 3.5.1 and table 3.8.1). However, the distribution of Grit change shows fatter tails, that is, has a lower peak, and thus more probability mass further away from the mode (figure 3.5.1). Moreover, Grit shows the highest percentage of individuals who change by at least two standard deviations (table 3.8.1). Table 3.8.1 also shows relatively high percentages of individuals with changes in extraversion and openness, both of which show no average within-person change. This finding shows that heterogeneous changes in a non-cognitive skill might exist even when, on average, this non-cognitive skill shows no development. This finding also shows that investigating only average changes might be insufficient for fully describing non-cognitive skill development. In sum, while all the non-cognitive skills we investigate appear to develop heterogeneously, Grit develops even more heterogeneously than the Big Five.

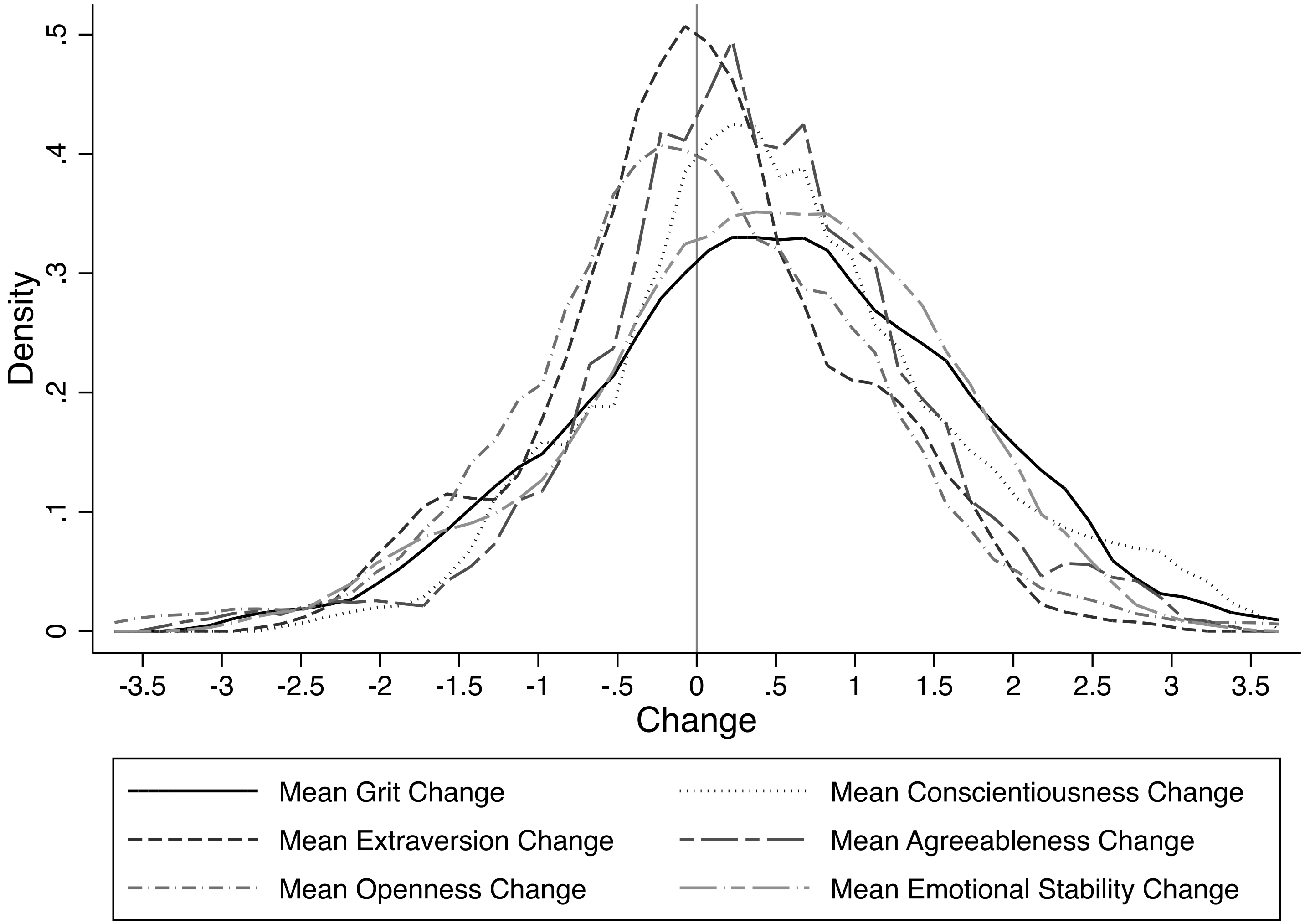

Figure 3.5.1: Density of Change Variables

Notes: All changes are standardized by the respective standard deviation of the baseline score. Leading House Apprenticeship Panel, Authors' calculations.

Next, we investigate the potential sources of this heterogeneity. The heterogeneity in the changes could result from either cumulative inputs before our baseline skill measurement or inputs between the two measurements.[6] Sections 3.5.1 and 3.5.2 examine both potential sources of heterogeneity: First, we study the relationship between baseline score and changes; second, we employ a traditional education production function to explain the heterogeneity that we observe in the data.

3.5.1 Heterogeneity over Baseline Scores

The first source of heterogeneity is captured by the baseline score of each individual. Figure 3.8.2 shows the change variable with respect to the baseline score. The changes are clearly related to the baseline scores: the higher the score, the lower the change. Likewise, the correlations in Table 3.3.1, column (3), show that the correlation between the baseline score and the change is much higher (in absolute terms) than the correlation between the final score and the change. Moreover, the correlations between the baseline score and the change are all very similar for Grit and the Big Five traits—except for extraversion—and range between -.56 and -.65 (table 3.3.1, column 3). These findings show that the development of a non-cognitive skill depends on the differences in its baseline score. However, these findings could be (a) a statistical artifact[7] or (b) an indication of decreasing marginal developments in non-cognitive skills. For the second explanation, our results imply that initial differences might level out over time and that individuals' non-cognitive skills tend to become more similar over time.

[6]This argumentation is in the spirit of a classic value-added model of education production (e.g., Hanushek, 2002).

[7]These findings might also be a typical example of *regression to the mean.* Statistically, high-baseline individuals are much more likely to end up at lower levels, and low-baseline individuals are much more likely to end up at higher levels. Therefore, the gap between the two groups might become smaller over time. Similarly, the observed relationship might also partly result from the way in which the measures are constructed, that is, they have a pre-specified range and thus are top-coded. In other words, the potential growth of an individual with high-baseline values is smaller than that of an individual with low-baseline ones. While we observe only a few top-coded observations in Figure 3.8.2, this top-coding problem could still exist separately for each different Likert-scale item but not necessarily for the sums of the different Likert-scale items that we use for the analysis.

We show that cumulative inputs before our baseline skill measurement—-captured by the baseline score—partly explain the heterogeneity in changes. Technically, our results thus show the importance of including the baseline score, when modeling the development of non-cognitive skills.[8] At the same time, however, the baseline scores do not explain all changes in non-cognitive skills, a finding indicating that experiences between the two measures might also be responsible for the heterogeneous changes. The literature suggests that education affects a wide range of outcomes, including test scores and non-cognitive skills (Cunha, Heckman, & Schennach, 2010). We now explore this second source of heterogeneity.

3.5.2 Education Production Function

The framework of our analysis of educational output considers a general production function[9] such as

$$S = f(B, R, L) \tag{3.5.1}$$

where S represents the array of non-cognitive skills as our output variables and f is a function of inputs. We consider the following inputs: a vector B of student background characteristics (family inputs and student characteristics), a vector R of education resources, and a vector L characterizing the learning environment. We summarize education resources (R) in two categories: educational content (C) and educational quality (Q). Given that we are interested in the change of non-cognitive skills (ΔS), we include in f the baseline level of non-cognitive skills (S_{t-1}). We can thus rewrite equation 3.5.1 as follows:

$$\Delta S = f(B, C, Q, L, S_{t-1}) \tag{3.5.2}$$

[8] For a discussion of different model specifications, see Allison (1990).

[9] For a detailed introduction of this type of analysis, see Todd and Wolpin (2003) or Hanushek (2002).

Assuming equation 3.5.2 to be additive separable, we can estimate it using OLS.[10]

Tables 3.8.2 and 3.8.3 present a detailed list and description of all the variables included in each vector of inputs, i.e., in C, Q, and L. To classify the inputs, we use these three reasonable categories—C, Q, and L—based on the classical education production function approach (e.g., Hanushek, 2002 and Todd & Wolpin, 2003) and further adjusted to the vocational education and training context. The first vector of inputs—educational content (C)—uses information on occupation and track, grades, and fixed training wage. The second vector of inputs—educational quality (Q)—uses measures based on teacher-student and trainer-student interaction, e.g., class size and firm size. The third vector—learning environment (L)—uses measures for intrinsic and extrinsic motivation, e.g., pay for performance and subjective student evaluations. The background variables include being a native speaker of German, mother's education, and the final middle school math grade.[11]

To better understand each input's contribution and to make our estimations feasible, we estimate equation 3.5.2 separately for C, Q, and L (always conditioning on B and S_{t-1}). The small sample size and the high number of inputs leaves us with limited degrees of freedom when estimating full models with all inputs. In addition, a model with such a high number of inputs would be over-controlled, and the residual effect of each input would be almost meaningless. Moreover, in our case, estimating the models separately for each single input generates essentially identical results to those we present next.[12] In sum, while we report the results for models using the three outlined categories of inputs (C, Q, and L) we ensure that our results do not depend on the categorization of the inputs.

Table 3.5.1 presents the results divided by inputs. For Grit, we find that the learning

[10] As a robustness check, we relax this assumption and also estimate models that allow for interaction effects between the inputs and S_{t-1}.

[11] As student's gender is highly correlated with occupation, one of our educational inputs, we do not include it as a background variable.

[12] Only one effect for emotional stability is not robust to this approach.

environment (L) explains part of the change (table 3.5.1, column 1). All of the significant inputs (*Satisfied with Prospects of Income, Job Meets Expectations, Over-challenged*) are subjective student evaluations, showing the importance of soft factors for the development of Grit. We find no effect for educational resources (C and Q).[13]

For the Big Five, none of the variables in Q, C, or L consistently explains the change (table 3.5.1, columns 2 to 6). Surprisingly, we find no effects for conscientiousness, which has substantial mean level changes and which the literature views as closely related to Grit (e.g., Duckworth et al., 2007). In contrast, the development of agreeableness and emotional stability can be explained in part by educational inputs. Moreover, having a school- or work-related conflict has a negative effect on the development of agreeableness and a positive effect on the development of openness. These contradictory effects show that certain educational inputs might affect the development of non-cognitive skills in opposite directions. Thus this finding supports the argument that higher levels of a personality trait are not necessarily always better (e.g., Almlund et al., 2011). In sum, the only variables that consistently remain statistically significant are the baseline scores, again showing the importance of the cumulative inputs before the baseline measure and underlining the importance of having a panel structure when studying the development of non-cognitive skills.

As we show the importance of the baseline score of a non-cognitive skill for its development, we also estimate models that allow the effect of the educational inputs (Q, C, L) to depend on this baseline score (S_{t-1}). Thus far, by simply controlling for the baseline score, we estimate average effects, i.e., we hold the baseline score constant for all individuals. Now, by interacting our inputs with the baseline score, we allow the effect of an input to differ for different initial levels of the respective non-cognitive skill. In this setting, the main effect of each input represents the effect for individuals with an average baseline score.[14] In addition, the main effect and the interaction effect taken

[13] As A robustness check, we use HC3 standard errors to correct for the limited sample size. All significant effects reported in Table 3.5.1 remain significant, except for the effects for emotional stability.

[14] The baseline scores are standardized to have a mean of zero and standard deviation of one.

together represent the effect for individuals with a baseline score above or below the mean. Therefore, this approach allows us to investigate the heterogeneous effects over the distribution of the baseline score, especially the effects at points in the distribution that are further away from the mean.

For Grit, while we again find no effects for educational quality (table 3.8.4), we find two interaction effects for educational content (table 3.8.5). Our results show that while a higher training wage has no effect for individuals with average Grit, it has a positive effect on the development of Grit for below-average Grit individuals and a negative effect on the development of Grit for above-average Grit individuals. The opposite relationship holds for the final grade, which again has no effect for the average-Grit individual. For the learning environment (table 3.8.6), we again find significant main effects and significant interaction effects. The negative effect of being over-challenged exists for average-Grit individuals and is more (less) pronounced for above-average (below-average) Grit individuals. In sum, from both modes, with and without interactions with the baseline scores, we find that the learning environment plays a central role in the development of Grit. These findings show the importance of subjective individual evaluations for the development of non-cognitive skills.

For the Big Five, the analyses that include interactions with the baseline scores (tables 3.8.4-3.8.6) provide four stylized results. First, the results for the educational content show several significant interaction effects for emotional stability. Second, small classes have an effect in this model: They are significantly related to decreases in extraversion and have a positive (negative) effect on the development of emotional stability for students with below-average (above-average) baseline emotional stability. Third, when comparing the *adjusted* R^2 of the models, we find that our models perform relative poorly in explaining changes in extraversion.[15] This finding is in line with our previous results, as we also find no substantial changes for extraversion and as the small changes are also

[15]The comparable low *adjusted* R^2 is present in both types of models, those that simply control for the baseline score (table 3.5.1) and those that additionally include interactions with the baseline (tables 3.8.4-3.8.6).

relatively low correlated with the baseline score. Fourth, the learning environment again appears highly important for the development of the Big Five (table 3.8.6). Nonetheless, in addition to subjective individual evaluations, for the Big Five, the variable *Pay for Performance* is another main driver for the effects of the learning environment. This finding shows the importance of extrinsic motivation for the development of non-cognitive skills.

In sum, our results point to the role of workplace-based factors—e.g., the fulfilment of job expectations, the satisfaction with prospects of income, or the perceived workload—as opposed to educational resources in the development of non-cognitive skills. Therefore, our results give a first indication for further studies. However, we still find only relatively few significant results and, therefore, reveal only limited systematic relationships. Given our sample size, one possible explanation for the limited significant relationships could be the lack of statistical power.[16] However, an alternative explanation is that traditional models of educational production functions with a strong emphasis on the role of educational resources do not perform well in explaining changes in non-cognitive skills. This explanation is in line with Cunha et al. (2010), who suggest that the technology producing non-cognitive skills is not necessarily the same as that producing cognitive ability. Moreover, our finding of the importance of the learning environment also supports this argument.

Finally, the heterogeneous development of non-cognitive skills could be partly—or, in the worst case, entirely—attributable to measurement error. Given our sample size, this possibility is a potential threat. Therefore, the next section tests the robustness of our results with respect to measurement error.

[16] To assess the power of our models, we perform a standard power analysis for our sample size (N=140). To do so, we use a typical power threshold of .8 and a share of variation explained by covariates of .25. When assuming that our educational inputs are random, we reveal with such a power analysis a minimum detectable effect size of .37 standard deviation units (at the 10 percent significance level). Thus our models might not able to detect any effects below .37. Given our observed average changes in non-cognitive skills and typical effect sizes in other studies investigating the formation of non-cognitive skills (e.g., Dahmann & Anger, 2014), this threshold seems relatively large and shows that missing power might be a serious issue in our data set.

Table 3.5.1: Relationship between Non-Cognitive Skills and Educational Inputs (OLS)

	Grit (z-score) (1)	Conscientiousness (z-score) (2)	Extraversion (z-score) (3)	Agreeableness (z-score) (4)	Openness (z-score) (5)	Emotional Stability (z-score) (6)
a) Educational Content [N=130]						
Occupation						
High-track Commercial	0.0219	-0.0906	0.0909	0.4731*	0.0221	0.1960
	(0.2749)	(0.2599)	(0.3236)	(0.2590)	(0.2589)	(0.2612)
Low-track Commercial	-0.0360	-0.0459	0.2778	-0.0928	-0.1782	0.1856
	(0.2579)	(0.2272)	(0.2932)	(0.2723)	(0.2897)	(0.2369)
Electrician	0.1908	0.1411	0.0809	-0.0830	0.1444	0.1632
	(0.2623)	(0.2746)	(0.3076)	(0.2837)	(0.2354)	(0.3017)
Log(training wage)	0.4152	0.5867	0.1234	0.0064	0.3469	-0.2590
	(0.5188)	(0.4163)	(0.4855)	(0.4685)	(0.5091)	(0.4138)
Final Grade	-0.1680	0.2832	-0.2052	-0.2497	-0.4169	0.4657*
	(0.3232)	(0.2783)	(0.2562)	(0.2927)	(0.2762)	(0.2717)
Adjusted R^2	0.318	0.378	0.208	0.382	0.324	0.311
b) Educational Quality [N=140]						
Small Class	-0.0892	-0.0380	-0.2560	-0.0452	0.0048	-0.1657
	(0.1476)	(0.1436)	(0.1629)	(0.1373)	(0.1420)	(0.1565)
Small Firm	0.2222	-0.0856	0.0808	-0.0693	0.0101	0.0280
	(0.1479)	(0.1385)	(0.1513)	(0.1370)	(0.1388)	(0.1444)
Conflict	0.2045	0.1021	0.0139	-0.4016***	0.2854**	0.0727
	(0.1494)	(0.1310)	(0.1628)	(0.1363)	(0.1437)	(0.1458)
Adjusted R^2	0.324	0.410	0.241	0.380	0.332	0.306
c) Learning Environment [N=138]						
Pay for Performance	-0.1538	-0.1183	-0.3012*	-0.1405	0.1290	-0.0688
	(0.1528)	(0.1490)	(0.1678)	(0.1561)	(0.1592)	(0.1649)
Satisfied with Prospects of Income	-0.3214**	0.0539	-0.0191	0.0044	-0.0667	-0.1472
	(0.1395)	(0.1425)	(0.1559)	(0.1566)	(0.1503)	(0.1574)
Job Meets Expectations	0.3835**	0.0855	0.0358	0.1304	0.0185	0.1232
	(0.1731)	(0.1608)	(0.1776)	(0.1569)	(0.1584)	(0.1609)
Over-challenged	-0.5856**	-0.5991	-0.0073	-0.5899	-0.3745	0.5474*
	(0.2352)	(0.4162)	(0.4907)	(0.4718)	(0.4300)	(0.3172)
Under-challenged	-0.2437	-0.0899	-0.0954	0.1525	0.3005	0.2126
	(0.2666)	(0.2047)	(0.2246)	(0.2046)	(0.2308)	(0.2847)
Adjusted R^2	0.333	0.397	0.240	0.348	0.307	0.327
Baseline Score	YES	YES	YES	YES	YES	YES
Background Variables	YES	YES	YES	YES	YES	YES

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Background variables include being a native speaker of German, having a highly educated mother, and the final middle school math grade. The base group for occupation is polymechanic (a specialized type of mechanic).
Leading House Apprenticeship Panel, Authors' calculations.

3.6 Measurement Error and Reliability of Changes

A major issue in investigating changes in skills over time is measurement error, particularly in small samples. First, one might think that the early score is correct, that the real score is stable, and that all changes over time are simply measurement error. Second and even worse, one might think that both scores are measured with error. Therefore, we need to know whether our results are purely driven by noise and how much of the change is actually reliable. To tackle measurement error issues, we use a "reliable change index," which accounts for reasonable levels of measurement error for the investigation of the average changes, and we add noise to Figure 3.5.1 (previously discussed) for the investigation of the heterogeneous changes.

To understand how much of the observed change is reliable, we turn to the clinical literature and estimate "reliable change indexes." Jacobson and Truax (1991) develope a "reliable change index" (RC) which corrects any observed change for a reliability coefficient (r).[17] If the scale produces only measurement error, r will be equal to 0, and no reliable changes would exist. In contrast, when the scale entails no measurement error at all, r will be equal to 1, and all changes would be reliable (table 3.6.1, column 1). We follow the approach by Jacobson and Truax (1991) and rely on usual reliability coefficients for our measures which range around .8 (John et al., 2008).

Calculating these "reliable change indexes," we find that for Grit 43 percent of the individuals change in a reliable way, either positively or negatively (table 3.6.1, column 2). Even in our most conservative specification ($r = .7$), at least 23 percent of the sample actually experience a reliable increase at the 10 percent significance level (table 3.6.1, column 3). We also find a substantial percentage of individuals with reliable increases for conscientiousness, agreeableness, and emotional stability (table 3.6.1, column 3).

[17] $RC = \frac{x_2 - x_1}{\sqrt{2(s_1\sqrt{1-r})^2}}$, with s_1 being the standard deviation of the baseline score. We test $RC = 0$ with a t-test.

Table 3.6.1: Reliable Change of Non-Cognitive Skills

	(1) *Observed*			(2) $r = .8$			(3) $r = .7$		
	Decrease (%)	No Change (%)	Increase (%)	Decrease (%)	Unreliable (%)	Increase (%)	Decrease (%)	Unreliable (%)	Increase (%)
Grit	30.07	7.19	62.74	11.11	56.86	32.03	7.19	69.93	22.88
Conscientiousness	24.18	14.38	61.44	5.23	69.93	24.84	1.96	79.08	18.95
Extraversion	39.22	18.30	42.48	11.11	77.12	11.76	11.11	77.12	11.76
Agreeableness	24.84	20.92	54.24	6.54	73.20	20.26	6.54	73.20	20.26
Openness	45.09	13.73	41.18	13.07	73.20	13.73	8.50	83.66	7.84
Emotional Stability	29.41	8.50	62.09	11.11	58.17	30.72	9.15	67.97	22.88

Notes: N=153. Reliable change on 10% level.

Leading House Apprenticeship Panel, Authors' calculations.

The intuition behind the RC index is that small changes in non-cognitive skills are likely measurement error and thus are unreliable. However, assuming that all small changes are unreliable and setting them to zero would have no strong effect on the heterogeneity of changes. Instead, we can obtain a graphical representation of the impact of measurement error on the full distribution of changes by simulating counterfactual distributions of changes under different levels of error in the data. To do so, we duplicate our initial data set 1,000 times and, in each duplication, randomly draw those individuals whose change is unreliable.[18] After repeating this replication exercise for different levels of error, we plot the new distributions of changes.

Figure 3.6.1 is a graphical representation of how the distribution of change in Grit varies with different levels of error in the data. The figure shows the changes in counterfactual measures, compared to the observed change. Figure 3.6.1 suggests that although increasing measurement error reduces the heterogeneity in changes, this heterogeneity does not completely disappear. Again, we find similar results for the development of conscientiousness, agreeableness, and emotional stability.

In sum, we conclude that measurement error only partially explains the heterogeneous development of Grit and the Big Five components conscientiousness, agreeableness, and emotional stability. These findings suggest that Grit and three of the Big Five (conscientiousness, agreeableness, and emotional stability) change during adolescence and that the change—although positive on average—appears heterogeneous across individuals.

[18] In practice, we implement the following two-step procedure. First, for a given proportion of the sample (i.e., $1 - r^2$), we substitute the observed values (initial and final value) with random draws from a uniform distribution between zero and one. Second, we bootstrap our initial data set 1,000 times and, in each replication, randomly select the individuals who receive measurement error. For the RC analysis, we set $r = .7$ and .8. As we are interested in investigating the heterogeneity, we perform this analysis in a mean-preserving manner.

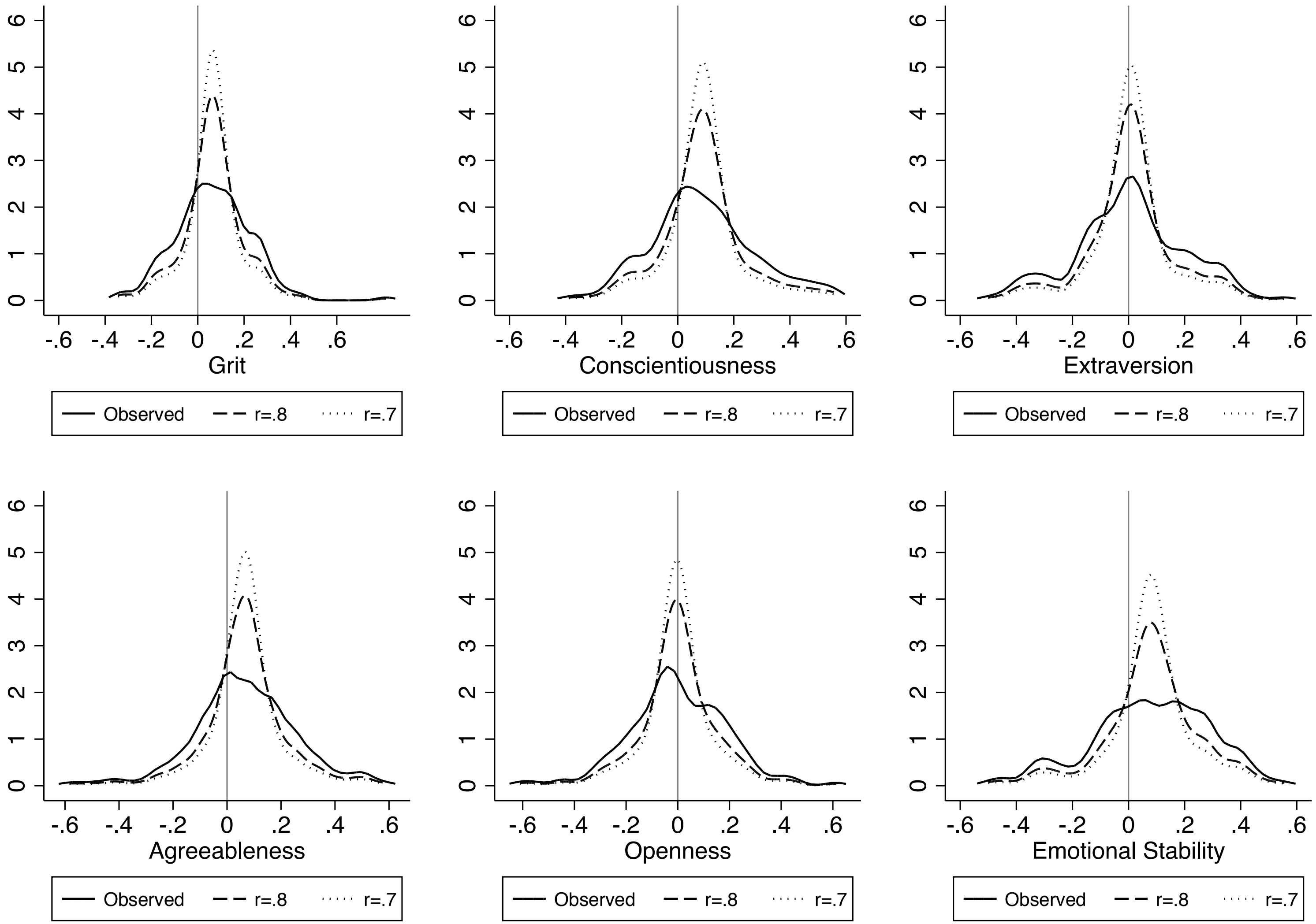

Figure 3.6.1: Density of Change Variables with Measurement Error

Notes: Kernel=gaussian, bandwidth=0.04.

Leading House Apprenticeship Panel, Authors' calculations.

3.7 Conclusion

Given the labor market relevance of non-cognitive skills, understanding both whether these skills change over time and, if so, how they are malleable is important. Our results show that a particular set of non-cognitive skills increases on average during adolescence: Grit and three of the Big Five personality traits (conscientiousness, agreeableness, and emotional stability) substantially improve on average between the ages of 15 and 22. However, these changes are heterogeneous, a finding that is of particular interest given our homogeneous sample in which all individuals received the same level of education. Moreover, we can rule out the possibility that all observed changes are merely measurement error by showing that these changes are robust to reasonable assumptions about measurement error.

In contrast to other studies, we find substantial changes in non-cognitive skills over time. Elkins et al. (2016) only investigate individuals enrolled in general education and conclude that personality (measured roughly by sets of adjectives) is stable for adolescents in school. For a similar age group, we investigate students enrolled in apprenticeship training to investigate whether work-based learning may have different effects. Thus our results suggest that work-based learning, with its more varied types of learning experiences, including its hands-on learning in a real business environment with coworkers and customer interactions, plays an important role in the formation of non-cognitive skills.

Using an educational production approach, we reveal an important role of workplace-based factors of the educational production process for the development of non-cognitive skills. Our findings show that subjective individual evaluations (e.g., the fulfilment of job expectations, the satisfaction with prospects of income, or the perceived workload) and—to a lesser extent—extrinsic motivation (i.e., pay for performance) are particular important for non-cognitive skill development and that educational resources might only play a minor role. While our results provide a first overview of the relevant variables for policy makers, future research needs to further investigate the specific relationships to derive more robust results that allow precise policy guidance.

Thus, while our results give first insights into the development of Grit during adolescence, they also leave some questions unanswered. For example, given our small sample size and limited variation in the training set-up, we are not able to fully explain the heterogeneity in changes. Therefore, further research investigating how different aspects of the learning environment in different types of education systematically affect the development of non-cognitive skills would be beneficial. Future research should examine these questions by using large-scale longitudinal data sets. Given the importance of Grit and given that we show that it changes over time, the inclusion of high-quality Grit questionnaires in large longitudinal data sets—such as the National Longitudinal Survey of Youth (NLSY), the Household, Income and Labor Dynamics in Australia (HILDA), the German Socio-Economic Panel (SOEP), or the National Educational Panel Study in Germany (NEPS)—would be particularly useful. Furthermore, these data sets would enable the investigation of the development of Grit during more diverse educational programs with potentially large differences in the learning environment.

3.8 Appendix

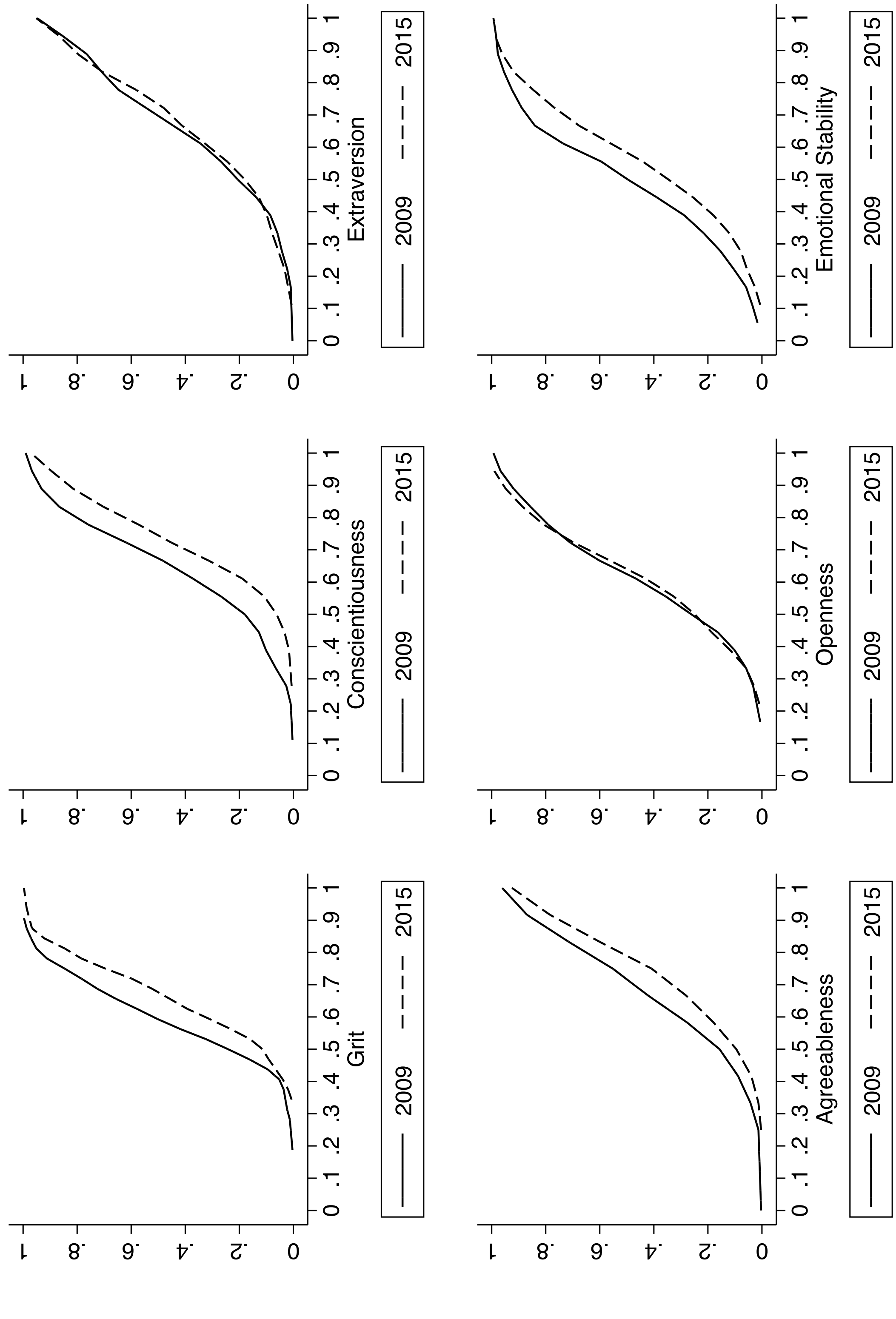

Figure 3.8.1: Cumulative Distribution Functions of Outcome Variables for Balanced Panel of Individuals

Notes: Leading House Apprenticeship Panel, Authors' calculations.

3.8.2 Change Variables

Table 3.8.1: Distribution of Change Variable

	$\Delta < \|1sd\|$ (%)	$\|1sd\| \leq \Delta \leq \|2sd\|$ (%)	$\Delta > \|2sd\|$ (%)
Grit	56.87	31.37	11.76
Conscientiousness	69.94	20.26	9.80
Extraversion	69.94	27.45	2.61
Agreeableness	73.21	18.95	7.84
Openness	73.20	20.92	5.88
Emotional Stability	58.17	36.60	5.23

Notes: N=153. Δ=change. sd=standard deviation.

Leading House Apprenticeship Panel, Authors' calculations.

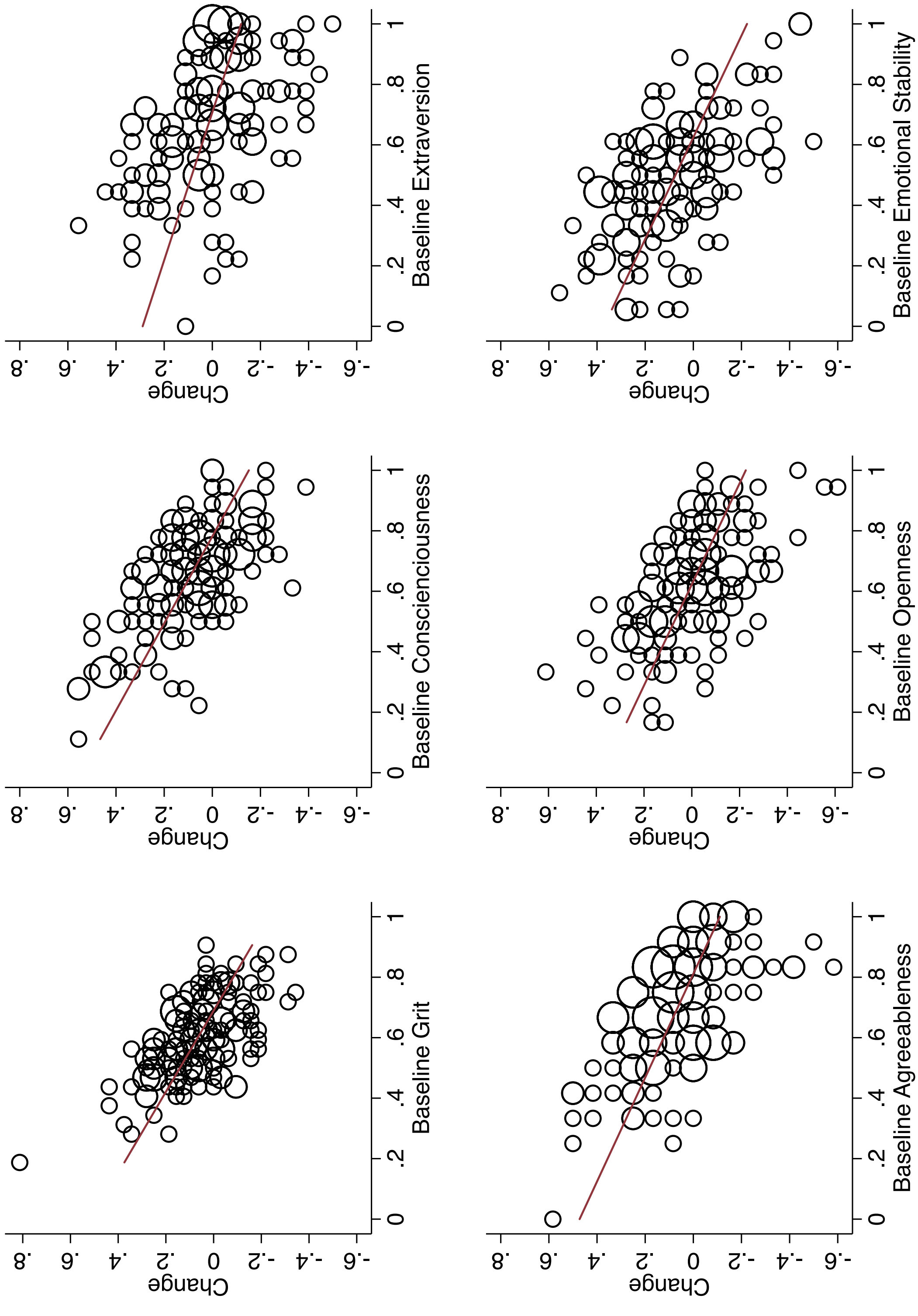

Figure 3.8.2: Change with Respect to Baseline

Notes: The size of the bubble indicates the number of identical scores.

Leading House Apprenticeship Panel, Authors' calculations.

3.8.3 Education Production Function Approach

Table 3.8.2: Description of Education Production Function Inputs

	Description
Background Variables (B)	
Native Speaker	=1, if native language is German
Mother Highly Educated	=1, if mother obtained a tertiary degree
Middle School Math Grade	Final math grade of middle school measured on a scale from 1 (worst) to 6 (best)
Educational Content (C)	
High-track Commercial	=1, if enrolled in commercial apprenticeship in profile M
Low-track Commercial	=1, if enrolled in commercial apprenticeship in profile B or profile E
Electrician	=1, if enrolled in electrician apprenticeship
Polymechanic	=1, if enrolled in polymechanic apprenticeship
Training Wage	Training wage measured in Swiss Francs per month
Final Grade	Final grade of apprenticeship training measured on a scale from 1 (worst) to 6 (best)
Educational Quality (Q)	
Small Class	=1, if class in vocational school has fewer than 20 students (median of class size = 20)
Small Firm	=1, if training firm has fewer than 100 employees (median of firm size = 100)
Conflict	=1, if indicated any conflict with master craftspeople, co-workers, teachers, or classmates
Learning Environment (L)	
Pay for Performance	=1, if receiving pay for performance based on performance in school, the firm, or both
Satisfied with Prospects of Income	=1, if answered that he/she is satisfied with the prospects of income in the job/occupation
Job Meets Expectations	=1, if indicated that the requirements of the job/occupation meet the expectations
Over-challenged	=1, if answered that he/she is over-challenged in the workplace or in school
Under-challenged	=1, if answered that he/she is under-challenged in the workplace or in school

Notes: Most variables are measured in the first or the second wave of the Leading House Apprenticeship Panel, i.e. in 2009 or 2010. Training wages are measured in 2011. The final grade of apprenticeship training is measured in the final survey in 2014/2015. Conflicts are measured at multiple points over time, i.e. in 2010, 2011, and 2012. Several variables are based on one or more 5-item Likert scale questions. Correspondence with a criterium is defined as scoring the two most extrem responses on these items. For the full questionnaires, see the additional material at the end of this dissertation, Bessey (2010), and Oswald (2013).

Table 3.8.3: Descriptive Statistics of Education Production Function Inputs

	N	mean	sd	min	max
Background Variables (B)					
Native Speaker	151	0.874	0.333	0.000	1.000
Mother Highly Educated	151	0.185	0.390	0.000	1.000
Middle School Math Grade	151	4.809	0.640	2.500	6.000
Educational Content (C)					
High-track Commercial	130	0.169	0.376	0.000	1.000
Low-track Commercial	130	0.469	0.501	0.000	1.000
Electrician	130	0.131	0.338	0.000	1.000
Polymechanic	130	0.231	0.423	0.000	1.000
Training Wage	130	1189.615	255.004	800.000	1650.000
Final Grade	130	4.759	0.274	4.100	5.400
Educational Quality (Q)					
Small Class	140	0.407	0.493	0.000	1.000
Small Firm	140	0.407	0.493	0.000	1.000
Conflict	140	0.643	0.481	0.000	1.000
Learning Environment (L)					
Pay for Performance	138	0.312	0.465	0.000	1.000
Satisfied with Prospects of Income	138	0.652	0.478	0.000	1.000
Job Meets Expectations	138	0.739	0.441	0.000	1.000
Over-challenged	138	0.036	0.188	0.000	1.000
Under-challenged	138	0.123	0.330	0.000	1.000

Notes: For a description of the variables, see Table 3.8.2
Leading House Apprenticeship Panel, Authors' calculations.

Table 3.8.4: Relationship between Non-Cognitive Skills and Educational Content with Interactions (OLS)

	Grit (z-score) (1)	Conscientiousness (z-score) (2)	Extraversion (z-score) (3)	Agreeableness (z-score) (4)	Openness (z-score) (5)	Emotional Stability (z-score) (6)
Baseline Score	2.5854	-1.2955	-2.0013	-2.6230	1.0277	1.2837
	(3.6150)	(3.1849)	(3.9935)	(3.6913)	(3.9323)	(3.1181)
Occupation						
High-track Commercial	-0.1149	-0.1248	0.0841	0.5422**	0.0171	-0.0330
	(0.2764)	(0.2962)	(0.3396)	(0.2716)	(0.2667)	(0.2437)
Baseline Score*High-track Commercial	0.0862	-0.0285	0.3041	0.0407	0.2538	0.7819***
	(0.2589)	(0.3451)	(0.2773)	(0.2382)	(0.3240)	(0.1963)
Low-track Commercial	-0.2369	-0.0763	0.2407	-0.0482	-0.2231	-0.0605
	(0.2579)	(0.2375)	(0.3227)	(0.2585)	(0.2919)	(0.2364)
Baseline Score*Low-track Commercial	0.2367	0.1082	-0.0276	0.2521	0.3370	0.7751***
	(0.2545)	(0.2431)	(0.2746)	(0.2447)	(0.3351)	(0.2141)
Electrician	0.1310	0.1156	-0.0262	-0.1598	0.1182	0.0079
	(0.2451)	(0.2828)	(0.2787)	(0.2833)	(0.2439)	(0.3124)
Baseline Score*Electrician	0.3067	0.0672	-0.6355**	-0.1561	-0.0469	0.4417**
	(0.2800)	(0.2836)	(0.2812)	(0.2623)	(0.3246)	(0.1922)
Log(training wage)	0.6267	0.5495	0.2564	-0.2244	0.5305	0.1078
	(0.5202)	(0.4472)	(0.5494)	(0.4610)	(0.4906)	(0.4343)
Baseline Score*Log(training wage)	-0.9532**	-0.1963	0.3000	0.0969	-0.2214	-0.5896
	(0.4621)	(0.4299)	(0.5240)	(0.4528)	(0.4817)	(0.3920)
Final Grade	-0.2518	0.2705	-0.1817	-0.3706	-0.4055	0.2821
	(0.2959)	(0.2884)	(0.2551)	(0.3008)	(0.2795)	(0.2719)
Baseline Score*Final Grade	0.7096***	0.4189	-0.1249	0.2610	-0.0503	0.3907*
	(0.2271)	(0.2850)	(0.2350)	(0.2604)	(0.2901)	(0.2156)
Adjusted R^2	0.366	0.369	0.237	0.384	0.312	0.358
Background Variables	YES	YES	YES	YES	YES	YES

Notes: N=130. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Background variables include being a native speaker of German, having a highly educated mother, and the final middle school math grade. The base group for occupation is polymechanic (a specialized type of mechanic). The baseline score is included as z-score.
Leading House Apprenticeship Panel, Authors' calculations.

Table 3.8.5: RELATIONSHIP BETWEEN NON-COGNITIVE SKILLS AND EDUCATIONAL QUALITY WITH INTERACTIONS (OLS)

	Grit (z-score) (1)	Conscientiousness (z-score) (2)	Extraversion (z-score) (3)	Agreeableness (z-score) (4)	Openness (z-score) (5)	Emotional Stability (z-score) (6)
Baseline Score	-0.6293***	-0.4475***	-0.3241**	-0.5654***	-0.3973***	-0.1456
	(0.1340)	(0.1283)	(0.1341)	(0.1329)	(0.1170)	(0.1101)
Small Class	-0.0939	-0.0425	-0.2815*	-0.0437	-0.0629	-0.1401
	(0.1492)	(0.1438)	(0.1633)	(0.1388)	(0.1493)	(0.1542)
Baseline Score*Small Class	-0.0061	-0.1513	-0.2025	-0.0753	-0.0563	-0.3563***
	(0.1439)	(0.1370)	(0.1933)	(0.1299)	(0.1536)	(0.1197)
Small Firm	0.2291	-0.0947	0.1005	-0.0668	-0.0050	-0.0084
	(0.1489)	(0.1400)	(0.1569)	(0.1385)	(0.1388)	(0.1434)
Baseline Score*Small Firm	0.0733	-0.1724	-0.1254	-0.1762	0.0209	-0.2113
	(0.1533)	(0.1390)	(0.1518)	(0.1214)	(0.1529)	(0.1388)
Conflict	0.2078	0.0861	-0.0103	-0.3916***	0.2704*	0.0690
	(0.1522)	(0.1394)	(0.1554)	(0.1393)	(0.1472)	(0.1408)
Baseline Score*Conflict	0.0637	-0.0739	-0.1457	0.0265	-0.2792*	-0.2376*
	(0.1538)	(0.1448)	(0.1649)	(0.1361)	(0.1502)	(0.1230)
Adjusted R^2	0.310	0.411	0.240	0.374	0.337	0.346
Background Variables	YES	YES	YES	YES	YES	YES

Notes: N=140. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Background variables include being a native speaker of German, having a highly educated mother, and the final middle school math grade. The baseline score is included as z-score.
Leading House Apprenticeship Panel, Authors' calculations.

Table 3.8.6: Relationship between Non-Cognitive Skills and Learning Environment with Interactions (OLS)

	Grit (z-score) (1)	Conscientiousness (z-score) (2)	Extraversion (z-score) (3)	Agreeableness (z-score) (4)	Openness (z-score) (5)	Emotional Stability (z-score) (6)
Baseline Score	-0.6151***	-0.7048***	-0.5278**	-0.6816***	-0.7102***	-0.4698***
	(0.1767)	(0.1513)	(0.2106)	(0.1614)	(0.2057)	(0.1394)
Pay for Performance	-0.1877	-0.0991	-0.3022*	-0.0777	0.0960	-0.0424
	(0.1566)	(0.1522)	(0.1731)	(0.1667)	(0.1603)	(0.1635)
Baseline Score*Pay for Performance	0.0983	0.2797*	0.0646	0.1110	-0.0865	-0.3763***
	(0.1815)	(0.1575)	(0.1725)	(0.1601)	(0.2048)	(0.1215)
Satisfied with Prospects of Income	-0.3204**	0.0212	-0.0123	0.0371	-0.0629	-0.1064
	(0.1386)	(0.1466)	(0.1618)	(0.1573)	(0.1583)	(0.1599)
Baseline Score*Satisfied with Prospects of Income	-0.0614	-0.0231	0.0293	0.2459*	0.2049	0.1275
	(0.1464)	(0.1398)	(0.1880)	(0.1362)	(0.1918)	(0.1450)
Job Meets Expectations	0.3841**	0.0284	0.0323	0.1688	0.0566	0.1649
	(0.1754)	(0.1646)	(0.1800)	(0.1594)	(0.1677)	(0.1587)
Baseline Score*Job Meets Expectations	0.1330	-0.0349	0.0100	-0.1171	0.0692	-0.0999
	(0.2229)	(0.1428)	(0.1947)	(0.1230)	(0.1690)	(0.1354)
Over-challenged	-0.7565***	-0.1253	0.0306	-1.1852***	-0.3656	0.8566***
	(0.1775)	(0.3940)	(0.8041)	(0.2440)	(0.4078)	(0.1997)
Baseline Score*Over-challenged	-0.3820**	0.7975***	-0.0610	-0.8344***	-0.1768	0.4559***
	(0.1630)	(0.2478)	(0.8820)	(0.2160)	(0.7635)	(0.1668)
Under-challenged	-0.2490	-0.0995	-0.1025	0.1554	0.3326	0.2080
	(0.2644)	(0.2203)	(0.2316)	(0.2164)	(0.2321)	(0.2785)
Baseline Score*Under-challenged	0.1221	-0.0877	-0.0610	0.0543	-0.0648	0.0180
	(0.2074)	(0.2041)	(0.1826)	(0.1449)	(0.2209)	(0.2514)
Adjusted R^2	0.322	0.408	0.210	0.373	0.289	0.352
Background Variables	YES	YES	YES	YES	YES	YES

Notes: N=138. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Background variables include being a native speaker of German, having a highly educated mother, and the final middle school math grade. The baseline score is included as z-score.

Leading House Apprenticeship Panel, Authors' calculations.

3.8.4 Attrition Analysis

Table 3.8.7: Attrition Analysis

	OLS	Probit	OLS	Probit
Grit 2009	-0.0043	-0.0110	-0.0015	-0.0035
	(0.0087)	(0.0218)	(0.0090)	(0.0228
Conscientiousness 2009	-0.0052	-0.0131	-0.0060	-0.0158
	(0.0117)	(0.0301)	(0.0121)	(0.0314)
Extraversion 2009	-0.0041	-0.0104	-0.0032	-0.0082
	(0.0093)	(0.0240)	(0.0093)	(0.0245)
Agreeableness 2009	0.0049	0.0125	-0.0006	-0.0011
	(0.0146)	(0.0377)	(0.0153)	(0.0385)
Openness 2009	0.0089	0.0228	0.0079	0.0206
	(0.0105)	(0.0263)	(0.0106)	(0.0267)
Emotional Stability 2009	-0.0026	-0.0069	-0.0055	-0.0146
	(0.0088)	(0.0223)	(0.0091)	(0.0228)
Background Variables	NO	NO	YES	YES
Observations	255	255	252	252
R^2	0.008		0.025	

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Dependent Variable: Being in the sample in 2015 (dummy variable). Background variables include being a native speaker of German, having a highly educated mother, and the final middle school math grade.

Leading House Apprenticeship Panel, Authors' calculations.

Table 3.8.8: Mean Differences between Non-Responders and Dropouts

	Dropout=0	Dropout=1	Difference	p-value
Grit 2009	0.631	0.548	0.082	0.018
Conscientiousness 2009	0.694	0.558	0.135	0.002
Extraversion 2009	0.698	0.697	0.001	0.990
Agreeableness 2009	0.705	0.658	0.047	0.308
Openness 2009	0.589	0.672	-0.083	0.074
Emotional Stability 2009	0.514	0.481	0.034	0.533
Native Speaker	0.768	0.800	-0.032	0.764
Mother Highly Educated	0.148	0.300	-0.152	0.114
Middle School Math Grade	4.839	4.689	0.150	0.317

Notes: N=102, with 82 "non-responders" and 20 "dropouts."

Leading House Apprenticeship Panel, Authors' calculations.

Chapter 4

Relative Importance of Personal Characteristics for Job Offers after Apprenticeship Training

Part of this chapter is a revised version of early parts of the working paper "The Relative Importance of Personal Characteristics for the Hiring of Young Workers" by Hoeschler and Backes-Gellner (2017).

4.1 Introduction

When making job offers, employers reveal their preferences about workers' personal characteristics (skills, abilities, and traits). However, employers face an uncertainty as to certain personal characteristics that might be unobservable. Especially non-cognitive skills, which a growing literature shows to be important in the labor market (e.g., Deming, 2017; Heckman & Kautz, 2012; Heckman et al., 2006), tend to be hard-to-observe for employers. This uncertainty might be even more pronounced for entry-level workers who have no employment history. To reduce this uncertainty, firms might screen entry-level workers' personal characteristics during programs such as internships, traineeships, or apprenticeships. For example, firms engaging in apprenticeship training learn about the ability of their trainees during the intense training period of several years (e.g., Acemoglu & Pischke, 1998). Afterwards, firms can decide to make job offers to these workers, thereby revealing their preferences for certain personal characteristics. We study these revealed preferences by investigating the influence of various personal characteristics on firms' job offers at the end of apprenticeship training.

A growing literature applies various research designs to investigate which personal characteristics employers value. This literature, which also suggests that non-cognitive skills are very important, uses five research designs. First, some studies use stated (discrete) choice experiments to elicit employer preferences over personal characteristics. In these studies, employers receive a set of CVs (or results of assessment centers) of potential applicants and are asked to select the candidates to whom they would offer a job (e.g., Biesma, Pavlova, van Merode, & Groot, 2007; Humburg & van der Velden, 2015). Second, correspondence studies attempt to identify the effect of the characteristics by exploiting real-world firm decisions. In these studies, researchers send out fake CVs with differing signals of personality to firms with job openings and observe the response rates (e.g., Protsch & Solga, 2015). Third, researchers simply ask employers for their preferences across a fixed set of personal characteristics (e.g., Biesma et al., 2007; Teijeiro, Rungo, & Freire, 2013). Such employer skill surveys are also frequently conducted by government agencies in various countries (e.g., CEDEFOP, 2014). Fourth, laboratory experiments study the effect of perceived personality on the likelihood of receiving a job (Baert & Decuypere, 2014). Fifth, to derive the skill demands of employers, some studies investigate job opening postings (e.g., Deming & Kahn, 2018). This approach is gaining popularity as such postings become more easily accessible through increasing online job searches.

However, all five approaches are limited in their ability to study employers preferences for at least two reasons. The first reason is related to firms' stated and revealed preferences: no study uses real-world high-stakes firm decisions that are relatively costly for the firms, and, therefore should reveal their true preferences. The second reason is related to firms' limited ability to observe important personal characteristics. To fully evaluate the relative importance of all personal characteristics, studies may use designs in which firms also know about the less-observable personal characteristics of individuals. Even though some studies have attempted to tackle this problem (e.g., Baert & Decuypere, 2014), none includes a substantial screening period in which employers can fully learn about the workers' characteristics. In contrast, our research design includes an extended screening period, thereby accounting for the limited observability of various personal characteristics, and uses job offers resulting in real employment.

Using this research design, we investigate the link between personal characteristics and job offers, both of which generally are hard to observe for researcher. We use a unique data set of Swiss apprentices, data that provide us with high-quality and well-established measures of intelligence, non-cognitive skills, and economic preferences—dimensions that many studies regard as unobservable ability. Moreover, we observe explicit job offers after an intense training and screening period. In our analysis we combine this information and compare the relative importance of different personal characteristics for the likelihood of receiving an offer. With our analysis we attempt to answer the following questions: Which personal characteristics are important for receiving a job offer at the end of apprenticeship training? Are these personal characteristics the same that make retained apprentices more likely to stay in the firm permanently? Do these offers matter for subsequent labor market outcomes?

We use jobs offers at the end of apprenticeship training because they should reveal firms' true preferences.[1] Towards the pre-specified end of the training, training firms can

[1]The transitions at the end of apprenticeship is intensively studied. These studies focus primarily on the analysis of wage differences between firm movers and firm stayers (e.g., Acemoglu & Pischke, 1998; Harhoff & Kane, 1997; B. Mueller & Schweri, 2015; von Wachter & Bender, 2006) and on factors affecting

make job offers and, if they do not, they face no firing costs because the training contracts simply run out. Moreover, in our sample, offers are selective and only 69 percent of the apprentices receive a job offer at the end of the training period. For the subsample that received an offer, 94 percent accepted. In addition, the distribution of the absolute values of the offers has a high mean and is very limited in range. Therefore, firms make job offers only to apprentices whom they really want to keep. In sum, these job offers are selective, result in employment at high wages, and thus are clearly not cheap talk.

In our main analysis we assess the power of different personal characteristics in explaining the likelihood to receive a job offer. Our type of analysis is related to Humphries and Kosse (2017), who investigate the importance of cognitive ability, personality, and economic preferences for high school GPA; to Burks et al. (2015), who investigate the importance of cognitive ability, personality, and economic preferences for college outcomes; and to Becker, Deckers, Dohmen, Falk, and Kosse (2012), who investigate the importance of economic preferences and personality for education, labor market and health outcomes. We add to these studies, which focus mainly on educational outcomes, by investigating firms' hiring decisions in detail.

We find that grades and non-cognitive skills are important for the likelihood of receiving a job offer. In contrast, we find no effects for intelligence or economic preferences. To investigate the relative importance of the different personal characteristics, we compare the predictive power of several personal characteristics. Only the models for grades, Grit, and the Big Five have predictive power in explaining offers. The relative predictive power of the Big Five is particularly striking, and they are by far the most important predictors. For the Big Five, we find that the baseline scores at the beginning of the training matter most, while for Grit we find that its development during training is most important.

Our results show that firms base their job offers on hard-to-observe non-cognitive

the probability of staying in the training firm (Euwals & Winkelmann, 2004; Franz & Zimmermann, 2002; Mohrenweiser, Wydra-Somaggio, & Zwick, 2017).

skills. Firms take this phenomenon into account by extensively offering programs that allow them to screen entry-level workers. Firms appear to use these programs to screen primary for non-cognitive skills. For policy makers our results imply that policies targeted towards the preparation of young people for entering the labor market might focus on improving non-cognitive skills because employers value non-cognitive skills as well as changes in them. In this regard, Chapter 2 and Chapter 3 provide valuable results as they show that non-cognitive skills are indeed changing during adolescence and adulthood. Moreover, policy makers may facilitate educational programs that include screening periods in which individuals can communicate their valuable non-cognitive skills to potential employers.

To show the importance of the investigated job offers, we relate them to later labor market outcomes. First, individuals who receive an offer at the end of the training period heavily reduce their search activities for a job outside the training firm. We observe significant and large differences in the number of job applications sent out and the number of months spent on job search. Second, we test whether the wages two years after training differ for individuals both with and without offers. Investigating wages two year after training, we find significantly higher full-time wages for those who received an offer. The wage difference between the two groups is equal to 621 Swiss Francs, or about 13 percent. In sum, we show that job offers are related to further labor market outcomes, and—as we show in our main analysis that the strongest predictors for receiving a job offer are the trainees' non-cognitive skills—provide an indication that differences in non-cognitive skills are valued by employers.

The link between offers—which are related to non-cognitive skills—and wages provides valuable insights to selection processes in the labor market. Studies which investigate wage differences between firm movers and firm stayers usually invest great efforts on establishing research designs that help to reduce the selection issues between the two groups (e.g., Acemoglu & Pischke, 1998; von Wachter & Bender, 2006). For example, while von Wachter and Bender (2006) find substantial wage differences between firm movers and stayers when not accounting for the selection into the two groups, they find no wage differences when estimating the causal effect of moving and staying. By describing this non-random selection process, our study reveals differences in non-cognitive skills between movers and stayers that provide an explanation for the different results in e.g., von Wachter and Bender (2006).

As we show that firms use non-cognitive skills when deciding to make retention offers, we also contribute to the growing literature on the importance of skills other than pure cognitive ability (e.g., Deming, 2017; Heckman et al., 2006). However, while the wage returns to non-cognitive skills are established in a causal manner (e.g., Heckman et al., 2006), in this paper we make no causal arguments, instead describing the usually unobserved selection processes into employment. As this is a major selection issue in labor economics, we think our study provides valuable insights for researchers, firms, and policy makers alike. To investigate this selection process, we use a unique research design that we describe extensively in the next section.

4.2 Research Design

Our research design has several favorable elements that enable us to tackle our research questions. These elements are based on our panel data that provides us with extensive measures of personal characteristics, and on the unique institutional setting of the Swiss apprenticeship training system. We will describe both aspects in this section.

4.2.1 Leading House Apprenticeship Panel

We use the Leading House Apprenticeship Panel, a panel data set which was started in 2009 with individuals who had just started their apprenticeship training in Zurich, Switzerland.[2] The training is conducted in three major occupations (commercial employee, electrician, and polymechanic) and—depending on the occupation—takes three to four years.[3] While the students receive one to two days per week of classroom learning in a vocational school, they receive most of their training in a host company (Wolter & Ryan, 2011). Given that the host company conducts such a substantial part of the training, it should be able to fully observe the trainee's personal characteristics.

We collected measures at the start of the apprenticeship training (at age 15-16), during the training, and two years after the training (at age 21-22). Figure 4.2.1 provides an overview of the time structure of the project. The training last three years for the commercial employees and four years for the technicians (electricians and polymechanics). For the main analysis, we use information provided in the initial and the final survey.[4] The initial survey provides us with rich baseline measures of personal characteristics and additional background variables. The final survey, which took place two years after the respective end of training, includes information on retention offers and further labor market outcomes.

Providing us with detailed information on job offers and personal characteristics, our data set is ideal to answer our research questions. First, our data set allows us to directly observe job offers, which register, firm-level, and large survey data sets rarely include. To this end, our data set provides detailed information on job offers by the training firm at the end of the training. By asking the trainees in great detail about their offers, we are confident that trainees truly report their offers.[5] Second, the data set provides us with measures for intelligence, economic preferences (time preferences and risk aversion),

[2]For a detailed description of the data and an overview of the entire project, see Chapter 3.3.1, Bessey (2010), and Oswald (2013).

[3]Vocational education and training is the main route of secondary education in Switzerland, serving 70 percent of young people (Hoffman & Schwartz, 2015). The occupations investigated in our study all rank in the top ten of the apprenticeship training occupations, with commercial employees outnumbering all other occupations by far (SERI, 2014).

[4]Additional surveys took place in 2010, 2011, 2012, and 2013 (only technical occupations).

[5]Ideally, we would like to match the trainee's information on the offers with additional information provided by the firms. However, this matching is not feasible as the Leading House Apprenticeship Panel surveys exclusively individuals and not firms.

and non-cognitive skills (Grit and the Big Five). Including such an extensive bundle of personal characteristics, our data set provides the unique opportunity to investigate the importance of all these characteristics for the likelihood to receive a job offer.

Another advantage of our data is that the surveyed population is very homogeneous with respect to occupation, education, and region. As all individuals received the same level of training, the restriction on range reduces selection issues. The setting we use is also similar to the reality of the hiring process, in which firms select workers for a given position within a given occupation. Also, studies investigating individual wage differences usually control for occupation and therefore investigate differences within occupations rather than between occupations (e.g., see von Wachter & Bender, 2006). Therefore, similar to Deming and Kahn (2018), we investigate the relative importance of various personal characteristics within narrowly defined occupations.

As with all panel data sets, we have to investigate attrition issues. The initial sample in 2009 consists of 265 individuals, 231 of whom provide measures of intelligence, grades, economic preferences, non-cognitive skills, and background variables. In the final wave, six years later, 159 individuals respond to our intense survey efforts (via e-mails, letters, phone calls, and social media), and 133 provide all analyzed measures. We view all our results as conditional on finishing apprenticeship training and staying in the sample. However, the overall attrition, which is 40 percent, is unrelated to intelligence, baseline non-cognitive skills, economic preferences, or various background variables (see table 4.6.5). For further details, we provide an in-depth attrition analysis for a similar sample in Chapter 3.3.3.

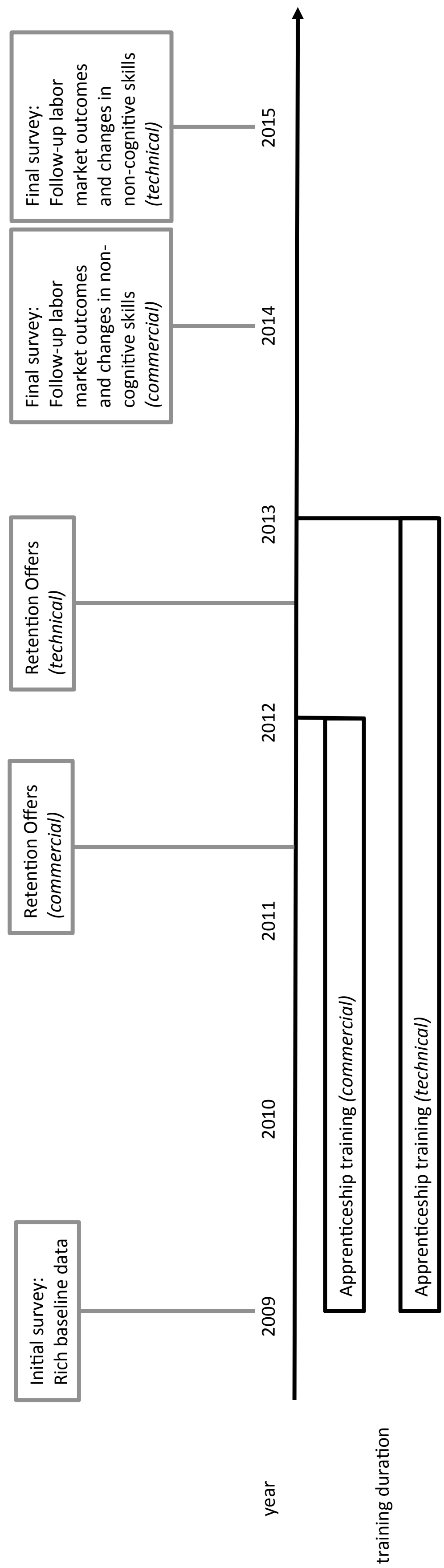

Figure 4.2.1: Time Structure

4.2.2 Measures of Personal Characteristics

To measure personal characteristics, we use well-established measures of intelligence, economic preferences (time preferences and risk aversion), and non-cognitive skills (Grit and the Big Five).[6] As our measures of intelligence, we use two tests: a general intelligence (IQ) test and the Cognitive Reflection Test (CRT). Our IQ test is one of the 11 modules of the Wechsler Adult Intelligence Scale (WAIS-III)—one of the most widely used IQ tests (Kaufman & Lichtenberger, 1998). We use the *digit symbol-coding test*, which asks subjects to match as many digits and symbols according to a given key as possible in a fixed time (for details, see Dohmen, Falk, Huffman, & Sunde, 2010). Our other measure of intelligence, the CRT, which is developed in Frederick (2005), asks subjects three questions, each having an intuitive answer that is incorrect. Finding the correct answer requires some cognitive reflection. However, once explained, the correct answer is easily understood. Both measures of intelligence are well established and appear to measure different facets of intelligence; the correlation between the two measures is basically zero.

Using these measures of intelligence, we show that the apprentices in our data set are average-ability students. The mean IQ score in our sample (table 4.2.1) is equal to the 50th-percentile score of a general sample of 16- to 17-year-olds in Austria, Germany, and Switzerland (von Aster, Neubauer, & Horn, 2009). Comparing the mean CRT score in our sample (0.81) to the scores of undergraduate students at selected public U.S. colleges (Frederick, 2005), we find that our score is in the range of the reported scores at the University of Michigan at Ann Arbor (1.18), Bowling Green State University (0.87), the University of Michigan at Dearborn (0.83), and Michigan State University (0.79). Moreover, Brañas-Garza, Kujal, and Lenkei (2015) survey 118 studies using the CRT and calculate for a total population of 44,558 students and non-students a mean of 1.19. Thus the apprentices in our data are clearly within an average-ability range.[7]

[6] For a detailed overview of the measures, see also Bessey (2010). Donata Bessey developed and conducted the initial survey in 2009. We are very grateful for her efforts which provide us with the opportunity to conduct our long-term analysis.

[7] This finding is a result of vocational education and training being the main route of secondary

For economic preferences, we use well-established paid experiments to measure patience and the willingness to take risks (Dohmen et al., 2010), both of which rely on choice tables.[8] Our measure of patience is the switching point X, at which individuals choose X today over 100 Swiss Francs in 3 month (for more details, see Oswald & Backes-Gellner, 2014). Our measure of the willingness to take risks is the certainty equivalent X, at which individuals choose a definite X over a coin toss yielding 5 Swiss Francs in expectation (for more details, see Bessey, 2010). Table 4.2.1 shows that our subjects are on average risk-loving. However, the modal certainty equivalent is equal to the expectation of the coin toss (5 Swiss Francs). The two measures of economic preferences are well-established and—as the two measures are uncorrelated—appear to measure different aspects of preferences.[9]

education in Switzerland, serving 70 percent of each cohort (Hoffman & Schwartz, 2015).

[8]Following the empirical findings of Meier and Sprenger (2015) and Andersen, Harrison, Lau, and Rutström (2008), we assume that economic preferences are stable over time.

[9]Burks et al. (2015) also find no correlation for similar measures of economic preferences in their data set of U.S. college students. For a more conceptual discussion on the relationship between time and risk preferences, see Andreoni and Sprenger (2012, 2015).

Table 4.2.1: Summary Statistics of Personal Characteristics

	Descriptive Statistics (1)				Correlation with Offer (2)
	mean	sd	min	max	coefficient
Intelligence					
IQ	75.940	13.028	45.000	116.000	0.0534
CRT	0.812	0.931	0.000	3.0000	0.0754
Grades					
Grade Middle School	4.891	0.396	3.667	5.667	-0.0474
Final Grade APT	4.783	0.267	4.200	5.400	0.1718**
Economic Preferences					
Willingness to Take Risks	5.857	1.587	1.000	10.000	0.0839
Patience	76.128	20.455	10.000	100.000	0.0809
Grit					
Grit_Initial	19.308	4.085	9.000	29.000	-0.0335
Grit_Delta	1.895	4.736	-11.000	12.000	0.1784**
Big Five					
Conscientiousness_Initial	11.579	3.182	2.000	18.000	0.1990**
Extraversion_Initial	12.429	3.964	0.000	18.000	-0.0306
Agreeableness_Initial	8.459	2.331	0.000	12.000	-0.1276
Openness_Initial	11.203	3.240	3.000	18.000	-0.1598*
Emotional Stability_Initial	8.820	3.590	1.000	18.000	-0.1293
Conscientiousness_Delta	1.699	3.364	-7.000	10.000	-0.1085
Extraversion_Delta	0.143	3.679	-9.000	10.000	0.0971
Agreeableness_Delta	0.797	2.120	-5.000	7.000	0.0823
Openness_Delta	-0.113	3.513	-11.000	11.000	0.1227
Emotional Stability_Delta	1.556	3.687	-8.000	10.000	0.1676*

Notes: N=133. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. IQ is the number of successively correctly matched symbols on the digit symbol-coding test of the WAIS-III. CRT is the number of correctly answered questions on the cognitive reflection test. The values for economic preferences indicate switching points in choice tables. The Grit measure is the sum of eight Likert scale items (0-4). The agreeableness measure is the sum of two Likert scale items (0-6). Each of the other Big Five measures is the sum of three Likert scale items (0-6). Emotional stability is calculated as the reverse of neuroticism (see page 43). The delta for each non-cognitive skill represents the difference between the respective final value and the initial value.

Leading House Apprenticeship Panel, Authors' calculations.

To derive our measures of non-cognitive skills, we use two well-established multiple-question inventories. For Grit, defined as the "perseverance and passion for long-term goals" (Duckworth et al., 2007, p. 1087), we use the 8-item Grit scale, a highly efficient and well-established questionnaire developed in Duckworth and Quinn (2009). Psychologists view this particular non-cognitive skill as one that is universally important in many domains and that has predictive validity over and above the Big Five personality traits (Duckworth et al., 2007; Duckworth & Quinn, 2009).

To measure the Big Five personality traits (conscientiousness, extraversion, agreeableness, openness, and emotional stability),[10] we use a well-established 3-item per trait scale based on the original Big Five Inventory (BFI) scale and further developed in Gerlitz and Schupp (2005).[11] The Big Five construct is the standard taxonomy for classifying personality traits (John et al., 2008).

We show in an accompanying investigation in Chapter 3, using the same data set of apprentices, that non-cognitive skills develop during apprenticeship training. Therefore, for all measures of non-cognitive skills, we include changes over time (*deltas*), which are the differences between the respective measure before and after the apprenticeship training (table 4.2.1). By construction each measure's delta is correlated with its initial value (for a discussion of this issue, see Chapter 3).[12]

In addition, our data set provides us with information on school grades, the easiest signals to observe for trainee characteristics.[13] We use two grades, measured at two points in time on the standard Swiss grade scale, which ranges from 1 (worst) to 6 (table 4.2.1). *Grade Middle School* is the average grade in Math, German, and English in the last year of full-time schooling before the apprenticeship training started. *Final Grade APT* is the final grade of the apprenticeship training after training. Measuring

[10]We calculate emotional stability as the reverse of neuroticism. For further details, see Chapter 3.3.2.
[11]For agreeableness, we can only use two items due to data issues.
[12]However, none of our results is affected by this correlation.
[13]For an overview of the relationship between grades, intelligence, and personality, see Borghans, Golsteyn, Heckman, and Humphries (2016). They show that grades, compared to IQ scores, are a better predictor for various important life outcomes because grades capture more relevant personality traits.

both education and training content, the final grade is based on grades (a) in vocational school and (b) hands-on tasks in training centers and the host company. Therefore, this grade measures general, vocational, and occupational skills. Moreover, our measures for intelligence and our measures for non-cognitive skills have both incremental predictive power for the final grade.[14] Therefore, the final grade constitute a credible signal for a certain set of personal characteristics, including cognitive ability, non-cognitive skills, as well as general, vocational, and occupational skills. As expected, the two grades are correlated ($r = 0.246$***).[15]

4.2.3 Job Offers

Our main outcome variable is a binary indicator for receiving a job offer after apprenticeship training. While doing their apprenticeship training, trainees are employed by host companies, where they receive a substantial share of their training. Therefore, firms can observe the trainees' productivity and screen for specific skills. In the final year of the training, the firm can decide to offer its trainee a permanent position after the training period is finished (for a detailed timeline, see figure 4.2.1). The trainee can decide to accept this offer or not.[16] The training period ends for all apprentices at the same time, which generates a spot market-like situation. All apprentices who are not retained by the training firm enter the secondhand market with all other un-retained apprentices (Ace-

[14]Following Borghans et al. (2016), who investigate the relationship between grades, intelligence, and personality, we can use our measures of intelligence (IQ and CRT) and non-cognitive skills (Grit and the Big Five) to explain the final grade. Therefore, we regress these measures individually and jointly on the final grade. Individual regressions show that intelligence (*adjusted* $R^2 = 0.0843$) performs much better than non-cognitive skills (*adjusted* $R^2 = 0.0399$) in explaining the final grade. A joint model (*adjusted* $R^2 = 0.1216$) shows that both, intelligence and non-cognitive skills, appear to be complements, and that—in comparison to the results of Borghans et al. (2016)—our models perform relatively poor in explaining the final grade. Thus the final grade might be highly influenced by other, uncorrelated skills, e.g., occupational and vocational skills.

[15]However, all the results we show for jointly estimated models, also hold for unconditional models.

[16]While each trainee clearly has only one training firm, a training firm could potentially have several trainees. By combining the observable information on firms to form unique cells, we find 80 combinations of firm characteristics, i.e., we observe at least 80 unique firms. Within one potential firm, the apprentices could still be in different departments with unrelated retention strategies. However, as we have no model for the interaction of several trainees in one firm with regard to the firm's retention offers, we assume a single-level model in which each firm employs one trainee.

moglu & Pischke, 1998) and theoretically become unemployed if they do not have other employment at the end of the training.

For job offers to reveal employers' true preferences, at least two requirements have to be fulfilled. First, employers must be unconstrained in their ability to make offers. Only when employers can freely decide to whom to make an offer, offers can reveal true preferences. Second, offers should be no "cheap talk." Therefore, making an offer needs to have real consequences for employers, i.e., in expectation, making an offer should lead to hiring the former trainee at a competitive wage. In the remainder of this section, we argue that in our case these two requirements are fulfilled.

Swiss firms are free in making job offers to their trainees at the end of the apprenticeship training. Given the low level of labor market regulation, firms are not constrained by institutional boundaries in their ability to make retention offers. Neither laws nor large-scale agreements between unions and firms cover the retention of apprentices. Moreover, firms can make offers to all their apprentices, even those with supposingly dominating outside options, such as other employment, further education, or even the mandatory military service. Actually, all of these options are to some degree endogenous to the offer, and none of these options restrains the firms' ability to make job offers to their apprentices. As long as the process of making an offer is relatively costless, firms can even make an offer to individuals who, they assume, would never accept it. In sum, firms are neither obligated to make offers to any of their apprentices nor restrained from making offers to all of them.

Another reason why Swiss firms are free in making offers is that they do not face any costs when they decide not to make one. On average training firms face no training costs from apprenticeship training (Muehlemann, Pfeifer, Walden, Wenzelmann, & Wolter, 2010)[17] and therefore do not need to retain a certain number of trainees to recoup

[17]Different occupations have different training costs. The technical occupations investigated in this study tend to lead to positive net training costs on average but with a high variance between firms (Strupler & Wolter, 2012). Moreover, the commercial apprenticeship training causes no substantial training costs on average (Strupler & Wolter, 2012). Therefore, we assume that for our full sample

such costs. Neither do firms face firing costs if they do not retain a trainee, as all training contracts just expire at the fixed end of the training. Not having any costs of separation is a major difference between our setting and up-or-out contracts, promotions in general, or other forms of retention, which may also provide settings for studying employer preferences. Our setting has the advantage of allowing firms to truly state their preferences without taking into account firing issues. In sum, firms are totally free in making offers.

Investigating the second requirement, we show that offers have a high likelihood to result in employment at high wages. In our sample, 69 percent of the apprentices receive a job offer at the end of the training period (table 4.2.2). Therefore, offers are selective, and—as not all trainees get one—we can use them to infer employer preferences. Figure 4.6.1 shows the distribution of retention wages which has a high mean, a clear lower threshold, and only limited variation.[18] Offers have a high mean wage of 4,642 Swiss Francs per month (or 55,704 Swiss Francs per year) with a low standard deviation of 393.[19] Such a limited variation in wages of young workers with the same level of education is in line with the literature on earnings dispersion, which shows that earnings fan out with workers' age (for an overview, see Neal & Rosen, 2000). Indeed, the limited variation in retention wages is the reason for our focus on explaining the likelihood to receive an offer and not on explaining the amount offered.

In the subsample that received an offer, 94 percent accepted (table 4.2.2).[20] This high acceptance rate shows that offers almost always result in employment. Moreover, it shows

positive training costs play no major role. We further investigate this issue in Chapter 4.4.1.

[18]Figure 4.6.1 shows a clear cut-off, as there are no offers below a certain threshold (3,500 Swiss Francs). Despite no general minimum wage in Switzerland, some sector-specific wage floors exist. In addition, firms appear to agree on an implicit lower bound for the wage offer. The observed distribution implies an equilibrium in which not every trainee simply receives an offer according to his or her marginal productivity but in which only the "best" apprentices receive offers and in which, therefore, offers have the potential to act as a credible signal for ability.

[19]The mean wage offer differs by occupations. Electricians have a statistically significant higher mean offer (4,965 Swiss Francs) than commercial employees (4,554 Swiss Francs) and polymechanics (4,691 Swiss Francs). The offered fixed pay per month is measured in intervals of 500 Swiss Francs. All reported additional payments are converted to monthly wages and added to the fixed pay. The limited variation might be partly due to the measuring of wages in relatively large intervals.

[20]The remaining 6 percent, who do not accept their offers, are not offered particular low retention wages.

that at this stage firms appear act as price setters, which can decide to make an offer or not. Afterwards, given they received an offer, almost all apprentices simply accept it. Put differently, firms appear to have some market power over the trainee.[21] The market power at the end of training could be based on various sources (for an overview, see Acemoglu & Pischke, 1999), at least two of which are related to the time elapsed since the end of training: low regional mobility[22] and asymmetric employer learning (Schönberg, 2007).

Therefore, we also investigate the likelihood to stay in the training firm more permanently, i.e., at least two years after training. In total, 55 percent of the apprentices who accepted the retention offer stayed in their training firm for at least two years (table 4.2.2).[23] We estimate all our models for both the likelihood to receive an offer and the likelihood to stay in the training firm for at least two years.

In sum, we conclude for two reasons that the offers should reveal employers' true preferences. First, firms can freely make offers. Second, the offered retention wage is high on average and varies little among trainees. Moreover, almost all trainees accept these offers, thereby forcing employers to pay these wages. Therefore, by investigating the likelihood to receive a job offer, we can observe employers' true preferences for certain personal characteristics.

[21]The market power might be monopsony power (e.g., Manning, 2011). Another reasons could be rent-sharing in a bilateral monopoly.

[22]Regional mobility in our sample is low but increases over time. Initially, almost all individuals in our sample live in the metropolitan area of Zurich. At the end of the apprenticeship training, 95.5 percent of them still live with their parents. Therefore, at the end of training regional mobility is very low. Two years after the training, 67.7 percent still live with their parents. Therefore, the early twenties appear to be a period in which individuals start moving out of their parents' places, a finding also observable in Swiss census data (FSO, 2016). However, at both points regional mobility is low.

[23]Our general transition patterns are in line with those of other studies for Switzerland. B. Mueller and Schweri (2015) find that 51 percent of apprentices stay with their training firm one year after training. Strupler and Wolter (2012) find a retention rate with the training firm of 37 percent during that first year. While a survey among graduates of apprenticeship training shows that 47 percent of graduates continue to work for their training firm (SERI, 2017), it shows differences in the retention rates in training occupations. In addition, the estimated retention rates crucially depend on the timing of the assessment. By definition the retention rate falls as the time between the assessment and the end of the training increases. Given our immediate assessment of the retention rate directly at the end of the training and the specific occupations we investigate, our estimated retention rates might be at the upper end. However, our data shows general transitions patterns that can help to explain the differences to other studies: relatively low mobility directly at the end of training coupled with high mobility within the first years after training.

Table 4.2.2: Descriptives of Transitions

Sample	N	Offer	Acceptance	Staying
Offer=1	92	100.0%	93.5%	51.1%
Acceptance=1	86	100.0%	100.0%	**54.7%**
Total	133	**69.2%**	64.7%	35.3%

Notes: Leading House Apprenticeship Panel, Authors' calculations.

4.3 Results

This section provides results for our main analysis, in which we investigate the relationship between personal characteristics and the likelihood to receive a job offer. Moreover, we compare the predictive power of our personal characteristics in order to asses which personal characteristics are most important for job offers. Further, this section provides results on the labor market relevance of offers and on the question whether moving or staying after receiving an offer is influenced by personal characteristics and whether such subsequent transitions are related to labor market outcomes.

4.3.1 Relationship between Personal Characteristics and Job Offers

To derive our results, we estimate OLS-models. First, we run individual regressions for each personal characteristic including a set of control variables. Second, to compare the relative predictive power of each personal characteristic, we compare the *adjusted* R^2s of models without control variables (for a similar approach, see Borghans et al., 2016). We use OLS in our main analysis as it provides us with a well-established measure of relative predictive power (*adjusted* R^2) that accounts for differing numbers of regressors.[24] Such a measure is crucial for our analysis, as our personal characteristics consist of differing numbers of variables. However, we show that all our results are robust to various other model specifications.

By providing the raw correlations between our personal characteristics and the likelihood to receive an offer, Table 4.2.1 shows already our main result that grades and various non-cognitive skills are important for receiving an offer. In detail, Table 4.2.1 shows that the final grade of the apprenticeship training is positively correlated with job offers. For the non-cognitive skills, the change in Grit and some of the Big Five variables are significant. Positive changes in Grit, higher initial level of conscientiousness, lower initial levels of openness, and positive changes in emotional stability are all correlated with the likelihood to receive an offer. However, for the Big Five the raw correlation of each trait is only of limited information, as we will primarily estimate joint models with all Big Five traits. When investigating these joint models, we have no theoretical expectation of the effect direction for each trait, and therefore, interpret the Big Five as a bundle and—in line with other studies (e.g., Becker et al., 2012)—do not discuss the effect direction of anyone of them. Indeed, each effect is simply a residual one conditioned on all the other traits. Given no clear procedure for attaching any meaning to these residual effects, we therefore treat all of the Big Five initial values as one variable and all the deltas as a second variable. Our main line of argumentation for all personal characteristics is then based on F-tests of joint significance and *adjusted* R^2s.

Moving towards the regression results, we find that the final grade of the training program is a significant predictor for receiving a job offer after training (table 4.3.1, column 2). Given that the training firm has three to four years to observe the trainee's abilities, the firm does not need to rely on grades as a signal for ability. In contrast, we

[24] The relationship between the R^2 and the *adjusted* R^2 (for degrees of freedom) is given as: *adjusted* $R^2 = 1 - \frac{(n-1)}{(n-K)}(1 - R^2)$, with n being the number of observations and K being the number of estimated coefficients. If the R^2 is sufficiently close to zero, i.e., when the sample correlation between the explanatory variables and the outcome is basically zero, the *adjusted* R^2 gets negative.

find no significant relationship between the likelihood to receive an offer and intelligence (measured by IQ and CRT).[25] These contradictory findings suggest that training firms value the final grade not because they constitute a measure of pure cognitive ability but because they measure occupational, vocational, and non-cognitive skills. As with our results for intelligence, we find no effects for economic preferences, i.e., for the willingness to take risks and for patience (table 4.3.1, column 3).

In contrast, we find that various non-cognitive skills have an impact on the likelihood to receive an offer (table 4.3.1, columns 4 and 5). For Grit, this likelihood is strongly related to the development of Grit during training. This result shows that employers value changes in non-cognitive skills.[26] It also links our findings back to Chapter 3 in which we show that Grit highly develops during apprenticeship training. For the Big Five, the likelihood to receive an offer is strongly related to the baseline personality at the beginning of training. However, all the changes taken as one group are not significant. An explanation for this result could be that training firms might be biased by the initial personality traits and might change their priors only slowly over time, an explanation that seems reasonable as the Big Five change less during apprenticeship training than Grit (see chapter 3). In sum, for the Big Five, the initial levels appear to be more important, while for Grit the changes over time appear to be more important. However, for both, the Big Five and Grit, we find that initial values and deltas are jointly significant, which is not surprising given the high correlation between the two measures (for a full discussion, see chapter 3.5.1).

[25] We also find no effect for CRT, when including four dummies, one for each potential outcome of the CRT, thereby allowing the CRT score to affect offers in a more flexible manner.

[26] Therefore, incorporating the development of non-cognitive skills over time when investigating the returns to non-cognitive skills is critical.

Table 4.3.1: Job Offers and Personal Characteristics (OLS)

	(1) Offer	(2) Offer	(3) Offer	(4) Offer	(5) Offer
IQ	0.0581				
	(0.0456)				
CRT	0.0023				
	(0.0429)				
Grade Middle School		-0.0360			
		(0.0362)			
Final Grade APT		0.0818*			
		(0.0441)			
Willingness to Take Risks			0.0376		
			(0.0360)		
Patience			0.0463		
			(0.0416)		
Grit_Initial				0.0397	
				(0.0461)	
Grit_Delta				0.1038**	
				(0.0486)	
Conscientiousness_Initial					0.1332**
					(0.0543)
Extraversion_Initial					0.0436
					(0.0521)
Agreeableness_Initial					-0.1407**
					(0.0557)
Openness_Initial					-0.1248**
					(0.0484)
Emotional Stability_Initial					-0.0991*
					(0.0546)
Conscientiousness_Delta					-0.0159
					(0.0528)
Extraversion_Delta					0.0376
					(0.0473)
Agreeableness_Delta					-0.0689
					(0.0500)
Openness_Delta					0.0122
					(0.0488)
Emotional Stability_Delta					0.0359
					(0.0485)
Controls	YES	YES	YES	YES	YES
F-test Joint	0.4115	0.1448	0.3469	0.0934	0.0002
F-test Joint Initial Values					0.0012
F-test Joint Deltas					0.6696
R-squared	0.060	0.074	0.061	0.083	0.211
N	133	133	133	133	133

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. All personal characteristics are standardized to have mean 0 and standard deviation 1. Controls include being a native speaker, mother's education, gender, and age. Listed is the p-value from a F-test for the joint significance of the indicated personal characteristics.

Leading House Apprenticeship Panel, Authors' calculations.

The findings in Table 4.3.1 are robust to various model specifications and estimation methods in the following four ways.[27] First, none of the effects depend on the inclusion of control variables (being a native speaker, mother's education, gender, and age), and the effects are virtually identical when we include controls for occupation. This shows that our results are similar across occupations and our personal characteristics do not just pick up differences across occupations. Second, the results do not depend on the applied grouping of personal characteristics. In Table 4.3.1 we include two variables for each personal characteristic (treating the Big Five as only two variables, initial values and deltas). However, this grouping has no effect on the results because the significant effects still remain significant in unconditional models with only one variable at a time. Third, when we estimate probit regressions, all effects remain significant. Given that our outcome is a binary variable, testing for such model specifications is crucial. Fourth, all effects remain significant when we use HC3 standard errors to correct for the limited sample size. Thus our overall conclusions do not depend on the particular model specification or estimation method and are robust to several other approaches.

Next, assessing the predictive power of the different characteristics, we find that the Big Five are the most important predictors for receiving a job offer. To display the relative importance of the different personal characteristics, Figure 4.3.1 shows *adjusted* R^2s for similar models as in Table 4.3.1 but without any control variables. Only the models for grades, Grit, and Big Five explain a significant share of the variance in the likelihood to receive an offer. To be more specific, grades and Grit perform about equally well, while the relative predictive power of the Big Five is striking, as it is about six times higher than that for grades or Grit. Running a full model with all personal characteristics (figure 4.3.1, column 6) shows that Grit, grades, and the Big Five all have incremental predictive power, as the bars of the separate models add up nicely in the full model. This finding shows that grades, Grit, and the Big Five appear to be complements.[28] A

[27] In some specifications, the effects for final grade and the Big Five (initial values) remain significant only at the 15 percent level. Specific results are available upon request.

[28] We test three additional specifications. First, including both measures that capture cognitive

simple variance decomposition for the model of Figure 4.3.1, column (6), again reveals that the Big Five, grades, and Grit are the important predictors. When we abstract from covariances, intelligence explains 2.2 percent of the explained variance, preferences explain 6.3 percent, Grit explains 11.1 percent, grades explain 16.6 percent, and the Big Five explain 63.8 percent. While these results could be partially due to the different number of variables for each personal characteristic, it again shows the importance of the Big Five.

To better understand the dominant effect of the Big Five, we now investigate the relative importance of the different Big Five traits. Therefore, we again perform a simple variance decomposition but now for the model of Figure 4.3.1, column (5). When we abstract from covariances, conscientiousness explains 33.1 percent of the explained variance, extraversion explains 0.9 percent, agreeableness explains 25.1 percent, openness explains 23.7 percent, and emotional stability explains 17.2 percent. These findings underline the dominant role of conscientiousness as the most important Big Five trait, a result that has been shown across many outcomes (Almlund et al., 2011). Moreover, our findings show that extraversion is not important for job offers, thereby adding to the general theme that not all Big Five traits are important for all outcomes (Almlund et al., 2011).[29]

We can compare our results for job offers to the results of studies assessing the relative importance of personal characteristics for educational outcomes. For example, using the same data set, Bessey (2010) shows that Grit and one Big Five trait (emotional stability) are related to the sureness to graduate from apprenticeship training, while she finds no significant relationships for intelligence, grades, and economic preferences.

ability—intelligence and grades—in one model, we get a model with relatively low predictive power (*adjusted* $R^2 = 0.0151$). Second, adding interactions between intelligence and grades to the aforementioned model, we find a small increase in the predictive power (*adjusted* $R^2 = 0.0328$) over the predictive power of the model that merely uses grades (*adjusted* $R^2 = 0.0233$). This finding might indicate that the power of intelligence may depend on grades, or vice versa. Third, running a model with both measures of non-cognitive skills—Grit and the Big Five—we get a predictive power (*adjusted* $R^2 = 0.1259$) that is about equal to the sum of the powers of the two separate models (i.e., one model for Grit and one for the Big Five). The last finding shows again that the two measures of non-cognitive skills appear to be complements.

[29]However, as personality might be valued differentially across occupations (Almlund et al., 2011) and sectors (Hamilton, Papageorge, & Pande, 2014), extraversion might be highly relevant in other settings.

However, she provides no F-tests for the joint significance of the Big Five which would be interesting to compare. Also, as she conducts her analysis at the beginning of the apprenticeship training, she includes no changes in non-cognitive skills. Burks et al. (2015) show that conscientiousness and—to a limited extend—patient time preferences are important for GPA and graduation on time in a sample of U.S. college students. They find no effect for intelligence when running a full model including several non-cognitive skills. Also, Borghans et al. (2016) show that non-cognitive skills predict test scores and grades above and beyond IQ scores. Therefore, the general patterns of our results, which explain the selection in the labor market, are in line with other studies using educational outcomes: non-cognitive skills appear to be the most important predictor across a variety of outcomes.

In sum, we show that the Big Five personality traits are by far the most predictive personal characteristic for explaining job offers. Moreover, we show a minor predictive role for Grit, another non-cognitive skills, and for grades, which are capturing a variety of skills including non-cognitive skills (for details, see chapter 4.2.2). In contrast, both, intelligence and economic preferences do not predict job offers. Therefore, we show that firms rely heavily on non-cognitive skills when making job offers after apprenticeship training.

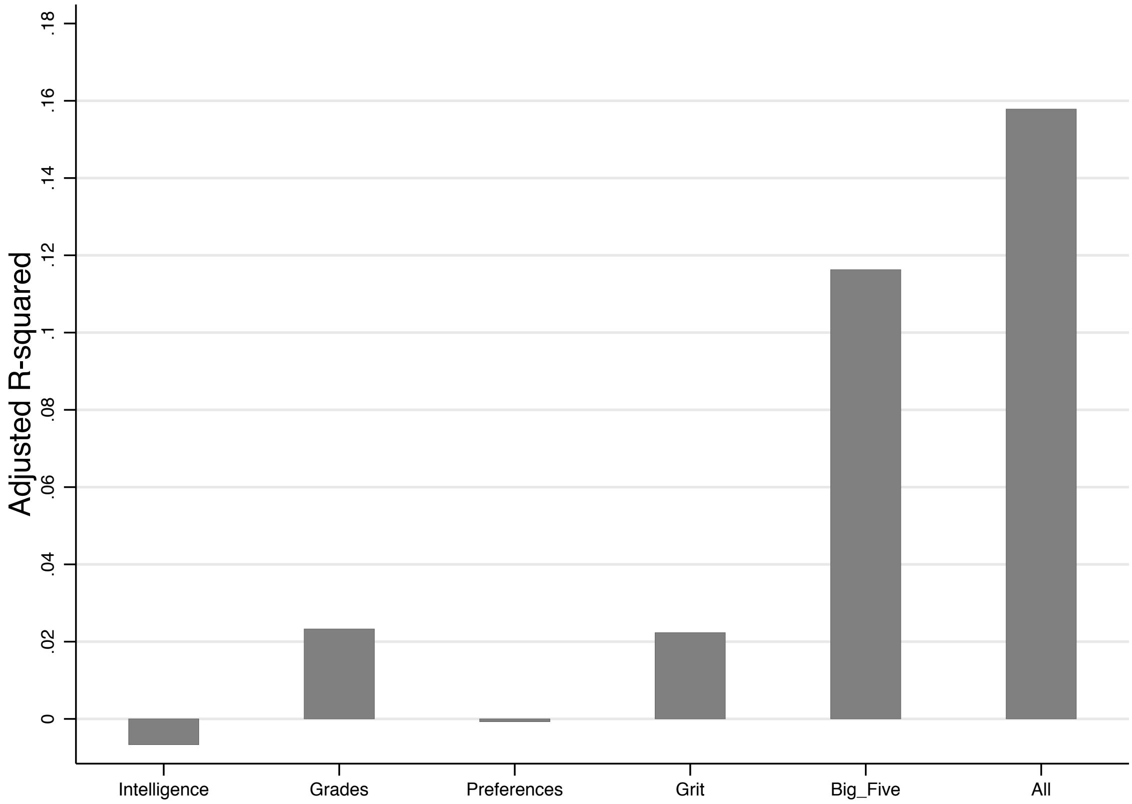

Figure 4.3.1: Relative Predictive Power for Job Offers

Notes: N=133. Column 1 to 5 show *adjusted* R^2 values of five individual regressions without control variables, one regression for each category of personal characteristics, i.e., one regression for each of the following: intelligence, grades, (economic) preferences, Grit, or Big Five (for a list of all variables included in each category, see table 4.2.1). For example, column 1 shows the *adjusted* R^2 of a regression of IQ and CRT on offer. Column 6 reports the *adjusted* R^2 of the full model (without control variables) including all measures of personal characteristics (for a full list, see table 4.2.1).

Leading House Apprenticeship Panel, Authors' calculations.

4.3.2 Offers and Labor Market Outcomes

Next, we analyze the question whether the investigated job offers matter for subsequent labor market outcomes. Therefore, we investigate two potential outcomes of receiving an offer: job search behavior directly after training and wages two years after training (table 4.3.2). First, individuals who receive an offer at the end of the training period invest significantly less in their search activities for a job outside of the training firm. We observe significant and large differences in the number of job applications sent out (1.1 versus 10.6) and the number of months spent for job search (0.4 versus 1.9).[30] Therefore, job offers appear to be a means for training firms to secure their market power directly after training because trainees who receive a job offer do not actively search for jobs outside of the training firm.

Second, we investigate the relationship between offers and wages two years after training. If only the "good" workers receive job offers, we should observe higher wages for workers who received an offer. Therefore, we test wether the wages two years after training differ for the workers with and without an offer. We find significant higher full-time wages two years after training for workers who received an offer.[31] The wage difference between the two groups is equal to 621 Swiss Francs, or about 13 percent (table 4.3.2). Therefore, the offers are highly important for the average trainee. While offers might also have a causal impact on wages, we view the wage differences primarily as a result of selection, and the offers as signal for personality characteristics. By investigating the group of retained apprentices in more detail, we will follow up on this issue in the next section. Therefore, we investigate what affects the mobility patterns of retained apprentices after training and whether these mobility patterns have an effect on wage differences within this group.

[30] As our survey provides no information on the timing of the offer and the job search activities, we can only describe the correlation and cannot show a causal effect of the offer on job search activities. To be precise, we cannot rule out the possibility that individuals might only get offers because they do not search for outside jobs.

[31] The percentage of individuals working full-time is relatively low two years after training (62 percent). The main reason indicated in our survey is "enrollment in further education and training." A high level of further education and training during the period following the initial apprenticeship training is a major characteristic of the Swiss education system (Hoffman & Schwartz, 2015; SERI, 2014).

Table 4.3.2: Labor Market Relevance of Offer

	Count	Offer=0	Offer=1	Difference	p-value
Job Search Behavior					
Search Time in Months	128	1.923	0.393	1.530	0.000
Number of Applications	128	10.641	1.135	9.506	0.000
Wages Two Years Later					
Wage	130	3278.312	4399.734	-1121.422	0.007
Wage (full-time equivalent)	130	3630.743	4821.502	-1190.759	0.007
Wage (only full-time employed)	80	4740.079	5361.441	-621.361	0.006

Notes: Monthly wages in Swiss Francs including any bonus payments or other additional yearly payments calculated on a monthly base.

Leading House Apprenticeship Panel, Authors' calculations.

4.3.3 Labor Market Mobility after Accepting Offer

In this section, we investigate the more permanent outcome of a job offer by explaining the likelihood to stay with the training firm for at least two years after accepting the offer. At this later stage, when various market forces start to work, the training firm has much less influence. First, the raiding activities of other firms and the training firm's interest in and ability to match outside offers become important (Lazear, 1986; Waldman, 1990).[32] Over time, offers become public knowledge, i.e., employer learning becomes symmetric (Schönberg, 2007). Individuals who received a job offer might use this offer as a credible signal during their subsequent career development. Given the institutional setting of the Swiss apprenticeship system, outside firms appear unable to directly observe the training firm's offer or to act on it by giving a counter offer. However, individuals who receive an offer and stay at the training firm reveal their offer by means of their employment patterns, thereby, allowing outside firms to observe it over time. Second, the apprentices'

[32] In line with Waldman (1990), we find that the wage for retained apprentices two years after training (mean: 4,517 Swiss Francs, sd: 2,076; full-time workers only: mean: 5,356 Swiss Francs, standard deviation: 941) varies much more than the retention wage offered (mean: 4,642 Swiss Francs, standard deviation: 393). However, this finding might be based on the survey design as both wages are measured in intervals of 500 Francs and these intervals appear to allow only for limited differentiation between the earlier (lower) wages.

preferences regarding employment may become more heterogeneous. While they might be primarily interested in securing any kind of employment directly at the end of training, they might later develop further interests, e.g., switching employers to obtain different sorts of working experiences or enrolling in further training. Therefore, we expect personal characteristics to be much less important at this later stage.

Conducting the same type of analysis as in section 4.3.1, we find that higher final grades are associated with staying in the training firm (table 4.6.1, column 2). Again, we find no effect for IQ or economic preferences. Moreover, we find basically no effect for Grit or the Big Five.[33] Given that non-cognitive skills are the major selection criterion in the first stage (receiving the job offer), the sub-sample of individuals who receives and accepts an offer, obviously varies much less in non-cognitive skills than our initial sample. This limited variation might explain the differing results. Moving towards a sub-sample replication of the comparison of powers, we again find that grades are the most important predictor for staying in the training firm at least two years after training (figure 4.6.2). No other personal characteristic has predictive power for explaining the likelihood to stay in the training firm. In contrast, the effect for grades in Figure 4.6.2 is increasing compared to the effect in Figure 4.3.1. Given that the most easily observable signal (grades) should become less important as workers stay longer in the labor market (e.g., Altonji & Pierret, 2001), this finding might be somewhat puzzling. However, the investigated workers are still young, and, thus firms might rely heavily on their grades because these constitute an easily observable signal also for external firms. Moreover, the final grade of the apprenticeship training also measures vocational and occupational skills, which might have an idiosyncratic value to the training firm.

Next, we show that firm movers and firm stayers, both of whom accepted the initial job offer and started to work for the training firm, do not differ in labor market outcomes (table 4.6.2). First, we find no significant differences in job search activities at the end of training—a finding that is not surprising, as both groups accepted the offer of the

[33] Only all Big Five variables, initial values, and changes are jointly significant (table 4.6.1, column 5).

training firm. However, this finding also shows that firm movers do not accept the initial offer simply because they could not find a better job. Indeed, neither of the two groups searches for jobs outside of the training firm at the end of training. Second, we investigate the wage differences between firm movers and firm stayers two years after training. Again, we focus only on full-time employed workers. Given that firm stayers by definition are still employed while firm movers could be anywhere, this restriction is crucial for comparing the two groups. In contrast to our findings in Chapter 4.3.2, we find no statistically significant wage differences between the two groups. In sum, when workers receive and accept an offer, whether they stay with the training firm or move to another firm within two years after the training is irrelevant for labor market outcomes.

4.4 Discussion and Robustness Checks

This section provides robustness checks that show additional results for the relationship between job offers and firm-related characteristics. Moreover, we discuss our results in more detail to address concerns regarding the generalizability of our findings and the reliability of the investigated intelligence measures.

4.4.1 Job Offers and Firm-Related Characteristics

In addition to trainee characteristics, firm-level and macro data could also affect the firms' retention decisions. In this subsection, we provide arguments for the limited confounding influence of these factors in our research design. In addition, we empirically test the relationship between several firm-related characteristics and the likelihood to receive an offer.

In our research design, firm-level and macro effects should not drive our results for two reasons. First, as our sample is very homogeneous and all firms operate in the same region, they all are exposed to the same general macroeconomic conditions, e.g., regional labor market thickness. Therefore, macroeconomic conditions should not affect our results. Second, while firms might use specific retention strategies unrelated to the trainee's personal characteristics, we argue that these types of strategies would clearly downward bias our results, i.e., make finding any significant effect unlikely. At the one extreme, a firm could always make each of its trainees an offer regardless of his or her individual characteristics, in which case trainees who received an offer should not have received one based on their characteristics. At the other extreme, a firm might never make an offer to any of its trainees, in which case some trainees who should have received an offer based on their individual characteristics do not. Both scenarios would decrease the difference in the mean characteristics between those who received an offer and those who do not, thereby causing our results to be downward biased (regression to the mean). Therefore, our estimates are only a lower bound for the importance of personal characteristics for firms' retention decisions.

Next, we empirically investigate the relationship between several firm-related characteristics and job offers. Table 4.6.3 provides descriptive statistics for the firm-related characteristics in our data set. When investigating retention at the firm-level, research shows that training costs affect retention at the firm level (e.g., Muehlemann et al., 2010). However, this is not necessarily true at the individual level (Muehlemann, Braendli, & Wolter, 2013). At the individual level, training costs might be endogenous and related to the trainees' personal characteristics (Muehlemann et al., 2013). If firms decide to provide the same level of training to all trainees, then higher-ability trainees might cause fewer training costs. However, firms might also provide more training to their higher-ability trainees in the expectation of retaining them, in which case higher-ability trainees might actually cause higher training costs. However, the extent of this strategy might be bounded by training regulations (for a complete discussion, see Muehlemann et al., 2013). Both levels of analysis—the firm- and the individual-level—identify firm size and industry/training occupation as two prominent factors affecting training costs. However, we find no statistically significant relationship between firm size or training occupation and the likelihood to receive an offer (table 4.6.4, columns 1 and 2).

Similarly, firms, that want to keep their trainees might already invest more in the trainee selection. If this is the case, firms which will keep their apprentices anyway, might just get better apprentices in the first place. One obvious way to attract "better" apprentices would be to pay higher training wages. However, we find no significant correlation between training wages and the likelihood to receive an offer (table 4.6.4, column 3). Finally, we investigate whether the likelihood to receive an offer depends on the interpersonal relationships of the trainee and his or her master craftsmen and co-workers. Again, we find no statistically significant relationship (table 4.6.4, column 4). In sum, we find no significant relationship between characteristics related to the training firm and the likelihood to receive an offer from this firm. Similarly, for the subsample that accepted an offer, we find no significant relationship between firm characteristics and the likelihood to stay in the training firm for at least two years (table 4.6.4, columns 5 to 8).

4.4.2 Initial Selection of Training Firms

Following up on the role of the firm in our setting, we further discuss the potential selection issue that arises as training firms choose their apprentices at the beginning of training and personal characteristics might already influence this selection process. Initially, students apply for apprenticeships with firms, firms then choose their apprentices from the pool of applicants. Ideally, to rule out the possibility that the characteristics that we find to be unimportant might actually be very important in the initial selection, we would like to replicate our study for the initial selection process. For such an investigation, we would need information on the personal characteristics of the full set of applicants, also the ones who received no apprenticeship position. Even if we have access to this data, such a design would not address the limited observability of non-cognitive skills and, therefore, would most likely yield very different results. Moreover, some of our results indicate that the initial selection process might be less important. First, we find no significant relationship between training wages and offers. Second, we find that changes in Grit during the training are important for receiving an offers.

Still, taking this potential limitation into account, we have to be clear that we address the question which skills are valued by employers conditional on having a pre-selected group of workers, i.e., trainees. This question is the typical one researchers ask when investigating all kinds of job promotions. However, when interpreting our results we need to be aware that all reported effects are conditioned on the first selection into the apprenticeships. Thus we can not rule out the possibility that economic preferences or intelligence might be highly predictive for entering into the apprenticeship training program.

4.4.3 Reliability of our Intelligence Measures

As we find no significant effects for intelligence, two concerns may arise regarding the reliability of our intelligence measures. First, the measures might be of low quality. However, our two measures are extensively used in the economic literature and proven to be useful (e.g., Dohmen et al., 2010). In addition, using our data set, we can empirically test the correlation between our intelligence measures and wages thereby investigating if these measures explain an important labor market outcome. While our IQ score (digit-symbol coding test of the WAIS-III) is unrelated to wages, the CRT score (cognitive reflection test) is significantly related to wages: increasing the CRT score by one standard deviation is associated with a 4.9 percent higher full-time wage two years after training. This supports our confidence that the CRT score measures important abilities. However, while a higher CRT is valued in the labor market, training firms do not take it into account when making their initial job offers at the end of training.

Second, from a conceptual perspective, our scales might measure other characteristics than intelligence. For example, our IQ measure might be better described as a measure for motivation as it is a relatively simple test, which was unincentivized (Almlund et al., 2011; Segal, 2012). Also, our CRT measure might asses several skills unrelated to intelligence (for a discussion, see Frederick, 2005). Purely following this interpretation of our measures, we would just find insignificant results for another set of (non-cognitive) skills, e.g., motivation. Put differently, not finding significant results for intelligence could also be due to our measures' limited ability to correctly assess intelligence. Acknowledging these considerations, we cannot rule out the possibility that more complex incentivized intelligence measures might lead to significant results.

4.5 Conclusion

We find that trainees' final grades and non-cognitive skills (Grit and the Big Five) predict job offers after apprenticeship training. These characteristics develop both before and during training. We show that the Big Five personality traits are the most important predictor. To show that job offers are a relevant outcome, we provide evidence for the labor market importance of the investigated offers: an offer is associated with fewer job search activities at the end of training and a substantially higher wage two years after training.

However, our results are limited in at least two ways that provide opportunities for future research. First, our sample size is small. On the one hand, smaller samples make finding significant results less likely. Therefore, finding significant effects in a small sample supports the robustness of our results. On the other hand, our significant results might gain even more credibility when replicated in larger samples for a larger set of occupations. Moreover, in larger samples, not finding statistically results for certain characteristics is a stronger argument for the minor role of these characteristics, because missing statistical power is less of a concern.[34] While our small sample size makes our insignificant results—especially for intelligence—less credible, it makes our significant results for the final grade,

[34]For a detailed power analysis for our data set, see Chapter 3.5.2.

Grit, and Big Five more credible. Second, our data provides no long-term labor market outcomes. Given the substantial percentage of part-time workers (38 percent) because many choose further training immediately after the initial apprenticeship training, future research should assess the labor market relevance of the offer in the longer term. Large longitudinal data sets with measures of personal characteristics, job offers, and wages would be necessary for overcoming the limitations of this paper.

By showing that firms use primary non-cognitive skills when making job offers after training, our results have implications for the literature on young workers. By describing the process of hiring decisions after apprenticeship training, we show that hiring after training is non-random (Gibbons & Katz, 1991) and that it is indeed best explained by differences in non-cognitive skills. In this regard, the results of our study show that accounting for the non-randomness of hiring is crucial when identifying the causal effect on wages of moving versus staying (e.g., von Wachter & Bender, 2006). However, we show that the worriesom selection is based on non-cognitive skills, not on cognitive ability, in line with recent research emphasizing the importance of skills other than cognitive ability (e.g., Deming, 2017; Heckman & Kautz, 2012).

Our results have implications for both firms and policy makers. We show that firms base their job offers after apprenticeship training primarily on hard-to-observe non-cognitive skills. One way how firms take this phenomenon into account is by extensively offering specific programs to entry-level workers, programs that allow them to screen these workers (e.g., internships, traineeships, or apprenticeships). Using these programs, firms primarily screen for non-cognitive skills. As another way to learn about applicants' non-cognitive skills, firms could also just deploy personality tests similar to the scales we use in this study. However, faking personality tests is very easy and we always need to consider test-takers incentives when interpreting such test results (for a general discussion, see Almlund et al., 2011). Especially under job applicants, faking personality tests happens frequently, and—with individual differences in the tendency to fake such tests—it heavily affects hiring decisions based on such tests (Rosse, Stecher, Miller, & Levin, 1998). In sum, our results show the importance of non-cognitive skills for firms' hiring decisions and—as personality tests are no convincing alternative—of extensive screening periods to learn about these hard-to-observe skills.

For policy makers, our study provides guidelines for the preparation of young people for the labor market. Indeed, once young people have attained a certain level of education, such efforts should focus on programs that target the formation of non-cognitive skills. In this regard, Chapter 3 provides evidence that non-cognitive skills are malleable during adolescence. Moreover, policy makers may facilitate training programs that include substantial screening periods and that thereby allow individuals to communicate their valuable non-cognitive skills to potential employers.

4.6 Appendix

4.6.1 Distribution of Absolute Offers in Swiss Francs

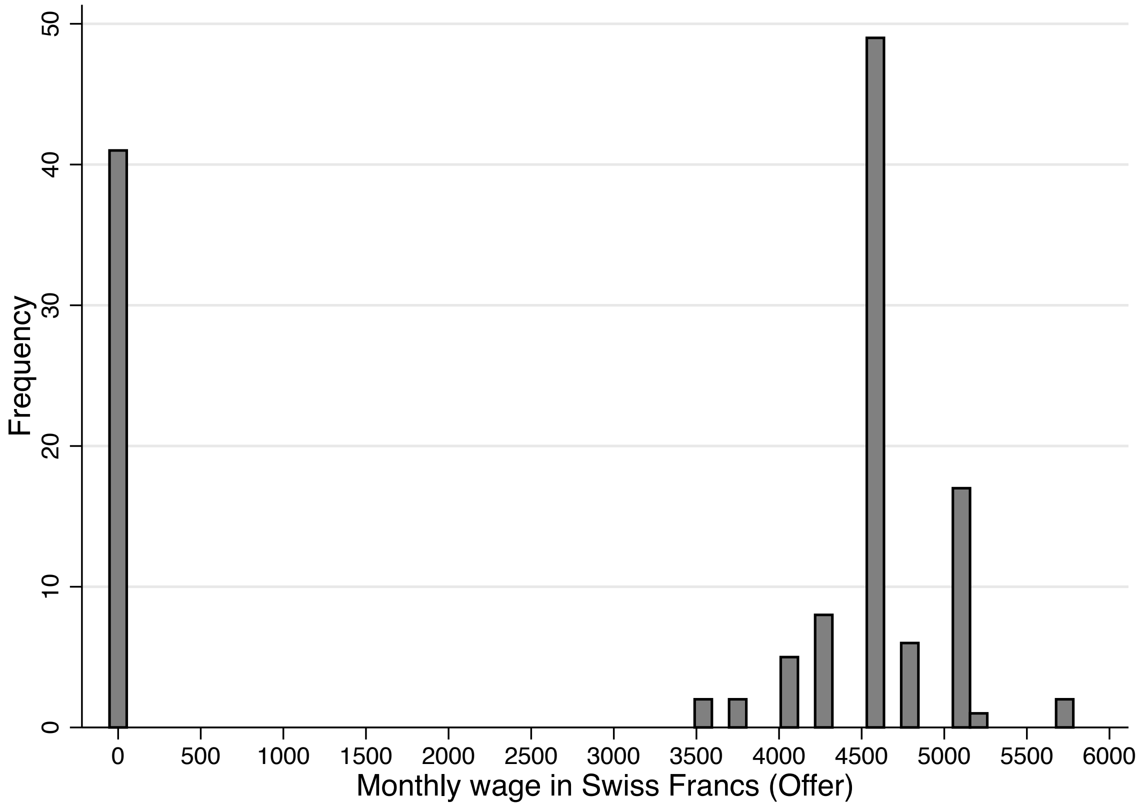

Figure 4.6.1: Distribution of Offers

Notes: N=133. The reported monthly wages may include additional yearly payments calculated on a monthly base.

Leading House Apprenticeship Panel, Authors' calculations.

4.6.2 Results for Staying

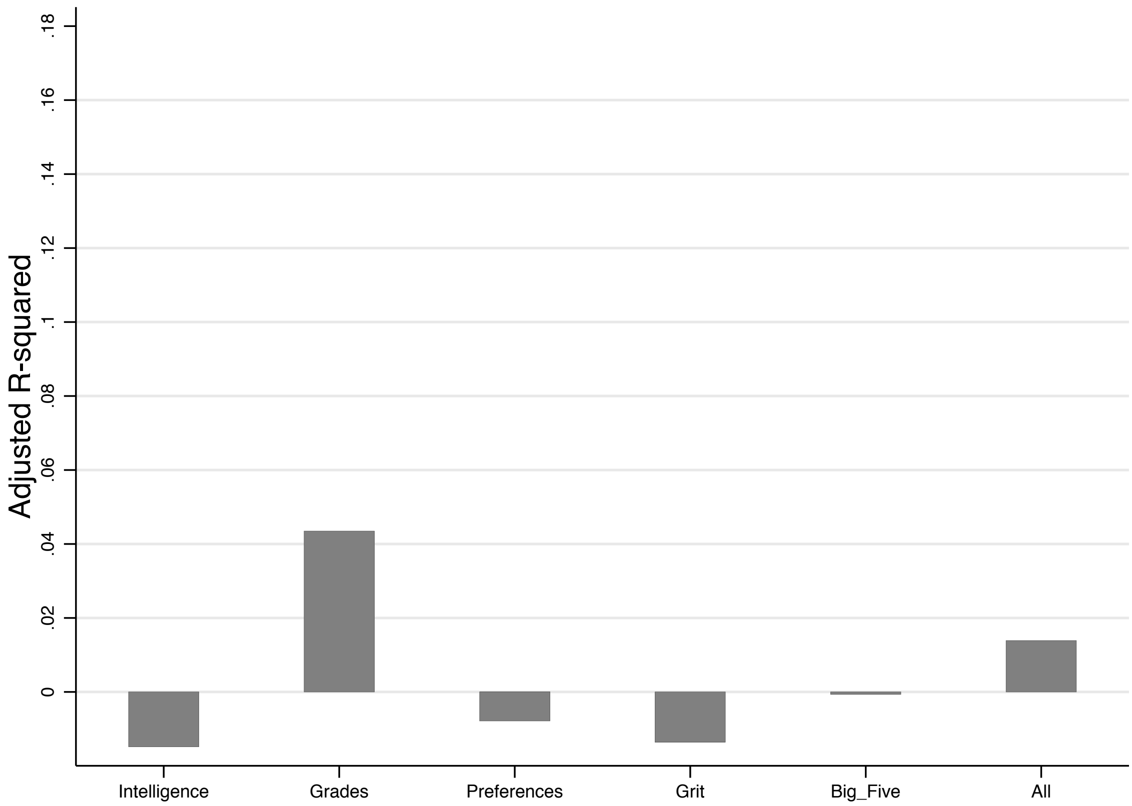

Figure 4.6.2: Relative Predictive Power for Staying - Retained Apprentices Only

Notes: N=133. Column 1 to 5 show *adjusted* R^2 values of five individual regressions without control variables, one regression for each category of personal characteristics, i.e., one regression for each of the following: intelligence, grades, (economic) preferences, Grit, or Big Five (for a list of all variables included in each category, see table 4.2.1). For example, column 1 shows the *adjusted* R^2 of a regression of IQ and CRT on offer. Column 6 reports the *adjusted* R^2 of the full model (without control variables) including all measures of personal characteristics (for a full list, see table 4.2.1).

Leading House Apprenticeship Panel, Authors' calculations.

Table 4.6.1: Staying and Personal Characteristics - Retained Apprentices only (OLS)

	(1) Staying	(2) Staying	(3) Staying	(4) Staying	(5) Staying
IQ	-0.0454				
	(0.0575)				
CRT	0.0310				
	(0.0578)				
Grade Middle School		0.0535			
		(0.0555)			
Final Grade APT		0.1218**			
		(0.0530)			
Willingness to Take Risks			-0.0400		
			(0.0508)		
Patience			0.0772		
			(0.0560)		
Grit_Initial				0.0349	
				(0.0669)	
Grit_Delta				0.0718	
				(0.0655)	
Conscientiousness_Initial					-0.0615
					(0.0806)
Extraversion_Initial					-0.0354
					(0.0720)
Agreeableness_Initial					0.1321
					(0.0911)
Openness_Initial					0.0565
					(0.0762)
Emotional Stability_Initial					-0.0095
					(0.0698)
Conscientiousness_Delta					0.0115
					(0.0657)
Extraversion_Delta					0.0787
					(0.0675)
Agreeableness_Delta					0.1568**
					(0.0714)
Openness_Delta					-0.0731
					(0.0667)
Emotional Stability_Delta					0.0200
					(0.0710)
Controls	YES	YES	YES	YES	YES
F-test Joint	0.6874	0.0123	0.1931	0.5466	0.0965
F-test Joint Initial Values					0.4921
F-test Joint Deltas					0.1676
R-squared	0.053	0.122	0.076	0.059	0.174
N	86	86	86	86	86

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. All trainee characteristics are standardized to have mean 0 and standard deviation 1. Controls include being a native speaker, mother's education, gender, and age. Listed is the p-value from a F-test for the joint significance of the indicated personal characteristics.

Leading House Apprenticeship Panel, Authors' calculations.

Table 4.6.2: Labor Market Relevance of Staying with the Training Firm - Retained Apprentices only

	Count	Staying=0	Staying=1	Difference	p-value
Job Search Behavior					
Search Time in Months	83	0.432	0.304	0.128	0.519
Number of Applications	83	1.027	0.674	0.353	0.485
Wages Two Years Later					
Wage	85	3955.329	4992.692	-1037.363	0.021
Wage (full-time equivalent)	85	4177.952	5450.873	-1272.921	0.008
Wage (only full-time employed)	58	5431.391	5294.417	136.974	0.586

Notes: Monthly wages in Swiss Francs including any bonus payments or other additional yearly payments calculated on a monthly base.

Leading House Apprenticeship Panel, Authors' calculations.

4.6.3 Results for Firm-Related Characteristics

Table 4.6.3: Summary of Firm-Related Characteristics

	Descriptive Statistics (1)					Correlation with Offer (2)
	count	mean	sd	min	max	coefficient
Small Firm	122	0.385	0.489	0.000	1.000	-0.0131
Occupation						
Electrician	133	0.105	0.308	0.000	1.000	0.1229
Polymechanic	133	0.241	0.429	0.000	1.000	0.0329
Commercial	133	0.654	0.477	0.000	1.000	-0.1089
Training Wage	120	1185.833	257.263	800.000	1650.000	0.0159
Conflict at Work	130	0.338	0.475	0.000	1.000	0.0308

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Small firms have below 100 employees (median of firm size = 100). Training wage is measured in Swiss Francs per month. Conflict at work indicates any conflict with master craftspeople or co-workers.

Leading House Apprenticeship Panel, Authors' calculations.

Table 4.6.4: Transitions and Firm-Related Characteristics (OLS)

	(1) Offer	(2) Offer	(3) Offer	(4) Offer	(5) Staying	(6) Staying	(7) Staying	(8) Staying
Small Firm	-0.0065				0.1836			
	(0.0851)				(0.1166)			
Polymechanic		-0.1337				-0.0312		
		(0.1265)				(0.1998)		
Commercial		-0.1334				0.1078		
		(0.1331)				(0.2096)		
Log(training wage)			0.2078				-0.2976	
			(0.2264)				(0.3159)	
Conflict at Work				-0.0024				0.0430
				(0.0874)				(0.1184)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
F-test Joint		0.5061				0.6859		
R-squared	0.049	0.053	0.055	0.044	0.071	0.053	0.069	0.047
N	122	133	120	130	79	86	78	86

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Controls include being a native speaker, mother's education, gender, and age. The base group in columns 3 and 6 is electrician. Listed is the p-value from a F-test for the joint significance of the occupation dummies.

Leading House Apprenticeship Panel, Authors' calculations.

4.6.4 Attrition Analysis

Table 4.6.5: Attrition Analysis

	OLS	Probit	OLS	Probit
IQ	0.0010	0.0010	-0.0288	-0.0774
	(0.0327)	(0.0858)	(0.0361)	(0.0950)
CRT	-0.0138	-0.0331	0.0020	0.0066
	(0.0345)	(0.0880)	(0.0369)	(0.0934)
Grade Middle School	0.0777**	0.2062**	0.0701**	0.1910**
	(0.0338)	(0.0898)	(0.0344)	(0.0910)
Willingness to Take Risks	0.0092	0.0276	0.0125	0.0356
	(0.0347)	(0.0872)	(0.0341)	(0.0883)
Patience	0.0322	0.0848	0.0221	0.0626
	(0.0341)	(0.0874)	(0.0346)	(0.0890)
Grit_Initial	-0.0040	-0.0094	-0.0071	-0.0158
	(0.0411)	(0.1038)	(0.0419)	(0.1068)
Conscientiousness_Initial	-0.0678*	-0.1801*	-0.0609	-0.1676
	(0.0401)	(0.1058)	(0.0396)	(0.1080)
Extraversion_Initial	0.0023	0.0094	0.0008	0.0047
	(0.0376)	(0.0961)	(0.0371)	(0.0981)
Agreeableness_Initial	0.0259	0.0672	0.0147	0.0376
	(0.0354)	(0.0914)	(0.0374)	(0.0934)
Openness_Initial	0.0052	0.0109	-0.0011	-0.0046
	(0.0367)	(0.0914)	(0.0364)	(0.0956)
Emotional Stability_Initial	-0.0138	-0.0388	-0.0011	-0.0046
	(0.0351)	(0.0894)	(0.0386)	(0.0983)
Controls	NO	NO	YES	YES
Observations	231	231	231	231
R-squared	0.036		0.064	

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Coefficients are indicated. Robust standard errors are in parentheses. Dependent Variable: Being in the sample in 2015 (dummy variable). Controls include being a native speaker, mother's education, gender, and age.

Leading House Apprenticeship Panel, Authors' calculations.

Chapter 5

Final Remarks

The main aim of this dissertation is to present a detailed analysis of the importance and malleability of different skills and personal characteristics during labor market transitions. I show that several non-cognitive skills are highly important for the entry into the labor market, and that these labor market-relevant skills are malleable over the life cycle.

The first study in this dissertation examines the development of self-esteem, one highly labor-market relevant non-cognitive skill, during adulthood. In this study, I investigate the relationship between unfulfilled educational aspirations and self-esteem. I use U.S. data spanning three decades to compare the self-esteem development of college dropouts with unfulfilled educational aspirations and that of graduates with fulfilled aspirations. To differentiate between the two groups, I develop an education classification that combines the highest type of college in which a student enrolled (aspiration) and the highest degree that he or she obtained. The analysis and its results make two major contributions. First, four-year college dropouts, when compared to graduates of four-year colleges, have permanently lower self-esteem. Second, finishing the highest type of college in which the student ever enrolled is critical for the formation of self-esteem. This contribution is derived from the finding that long-term negative effects exist for four-year college dropouts who obtain an associate's degree but not for associate degree holders who had never enrolled in a four-year college. In sum, this study shows that non-cognitive skills develop in adulthood but that their development depends on the fulfillment of educational aspirations.

The contributions of my first study imply that students should be supported in forming realistic aspirations. Moreover, by highlighting the role of community colleges for the formation of self-esteem, my study shows that—besides four-year colleges—other types of colleges might be beneficial for the formation of non-cognitive skills. This finding mirrors the results for the labor market returns of different types of colleges, which show that graduating from a community college outperforms merely attending a four-year college without obtaining a bachelor's degree (Kane & Rouse, 1995). Based on these results, I argue that even more diverse educational systems might be beneficial for the development of non-cognitive skills. Thus I test in my second study whether vocational education and training might provide such an avenue.

The second study uses a unique panel data set of Swiss apprentices to examine the malleability of non-cognitive skills during adolescence. In sum, the results of this study entail at least three major contributions to the literature on the development of non-cognitive skills. First, Grit and three of the Big Five personality traits (conscientiousness, agreeableness, and emotional stability) increase on average by about .5 standard deviation units between ages 16 and 22. Second, these changes are highly heterogeneous across my homogeneous sample. When explaining this heterogeneity in changes, I show—in addition to a prominent role of the baseline score—that workplace-based factors, e.g., the fulfilment of job expectations, the satisfaction with prospects of income, or the perceived workload, play a crucial role. Third, these average and heterogeneous changes are robust to reasonable levels of measurement error. By comparing these contributions with findings that show the stability of non-cognitive skills for a similar age group enrolled in general education (Elkins et al., 2016), I argue that apprenticeship training might be an effective intervention for fostering the development of non-cognitive skills. Given positive changes in Grit and the Big Five, I address the question whether these changes are valued by employers in the next study.

The third study in this dissertation examines the relative importance of several personal characteristics (including cognitive ability and non-cognitive skills) for job offers. I use the same data set as in the second study but now focus on describing the labor market transition of entry-level workers. I consider a setting that allows firms to observe potential workers during a long screening period (three- to four-year apprenticeship training) and thus observe non-cognitive skills, which are usually only limitedly observable during the regular hiring process. At the end of the screening period, firms can decide to make job offers to certain workers, thereby revealing their preferences about workers' personal characteristics. I am able to connect real-world job offers and workers' personal characteristics, both of which are usually unobserved by researchers.

When I apply this research design, the findings in this study make three major contributions to the literature on the relative importance of different personal characteristics. First, the final grade of the apprenticeship training is important for receiving a job offer. Second, non-cognitive skills are highly important, with the Big Five personality traits by far the most important predictor. Moreover, for Grit, its development during the apprenticeship training is most important. Third, neither intelligence nor economic preferences are important to firms in their hiring decisions in this context. These results underscore the importance of non-cognitive skills in the labor market. In addition, these results show an explanation driving this importance: that firms rely on non-cognitive skills in their hiring decisions.

The generalizability of the results in this dissertation is limited by two main boundaries: time and place. For issues of time, the question arises: How can results obtained from data from the past still be useful for current policy? While the second and the third studies use very recent data, study one uses data from as early as 1979 and from individuals born as early as 1957. However, as this study focus on long-term effects, I have to use data from a long time in the past and from relatively "old" cohorts. Moreover, although U.S. college enrollment has increased dramatically over the last 30 years, the dropout rates have remained relatively stable at around 55 percent (Bound & Turner, 2011). Thus I remain confident that the results and patterns are both informative and useful for policy makers.

The second boundary is spatial: What can other countries learn from the results of this dissertation? This issue particularly affects the policy implications based upon Swiss data on apprentices. On the one hand, Switzerland's apprenticeship training allows scholars to answer research questions that cannot be resolved by studying other countries. On the other hand, the results can only be generalized to countries with similar apprenticeship training. While some European countries have similar systems in place, the UK, the U.S., and various other countries are increasing their efforts to develop similar vocational training regimes. Thus my results based on Swiss data are of the utmost relevance and interest to educational policy makers worldwide who are interested in implementing such training systems. Moreover, the results from the third study should clearly generalize to all kinds of on-the-job training programs for entry-level workers, mainly internships and traineeships. More broadly, my results provide evidence for employer preferences in general.

Another remark concerns the causality of the results in this dissertation. The first study can be interpreted as a causal analysis under relatively strong assumptions. However, given the complexity of post-secondary education decisions, I cannot use random variation to identify my effects. Therefore, I have to rely on observable, potentially non-random, behavior, and have to exploit the panel structure of my data set to increase the credibility of my results. The second study is a descriptive analysis of an important growth process, that the literature has thus far overlooked. The third study identifies the variables that are related to an important real-world selection process. The explicit goal of this study is to describe a very important but usually unobserved selection process by employers. In this study, I am interested in describing the revealed preferences of employers after a long screening period in which the employers are able to learn about a variety of personal characteristics and in which personal characteristics might even change. Thus my primarily descriptive analyses indeed constitute the appropriate method for me to have applied.

The empirical findings and their policy implications described in this dissertation point to two avenues for future research. First, this thesis identifies the need for large panel data sets with relevant well-established measures for cognitive ability and non-cognitive skills. In this regard, my findings show that having such data is a relevant issue as non-cognitive skills change over time. These data sets should also include detailed information on individuals' labor market outcomes as my third study indicates that these outcomes are related to non-cognitive skills. Using such a data set, further research would be able to tackle some of the limitations outlined in these three studies. For example, such a data set would allow the investigation of the development and the labor market-relevance of different non-cognitive skills in more heterogeneous training and education set-ups. Second, further research should attempt to identify effective interventions targeted at the development of the non-cognitive skills that this dissertation has investigated. Such new research may focus on using randomized control trials or field experiments that randomly alter elements of the learning environment because my second study shows that soft factors of education might play a crucial role for the development of non-cognitive skills. Whenever such efforts result in designing new programs or interventions, the efficiency of such programs needs to be evaluated in the light of my findings. The question arises whether explicitly targeting non-cognitive skills generates positive returns and to what extent existing education programs might already produce these skills, something I show in my second study to be the case for apprenticeship training.

This dissertation provides an extensive overview of the development of labor market-relevant non-cognitive skills over different periods of the life cycle and in different types of education. It investigates the development of such skills both during adolescence and adulthood, and shows how both general education and vocational education and training might affect this development. Finally, it investigates how initial levels and changes in these skills might affect labor market transitions, thereby providing an argument for the labor market-relevance of these non-cognitive skills.

Taken together, the chapters of this dissertation clearly indicate at least three important policy considerations. First, policy makers are well advised to build a diverse educational system with many different paths and a high permeability between these paths. Such a system can enable all individuals to form realistic educational aspirations and to attain them. In this regard, the establishment and development of vocational education and training, which offers an additional path to recognized degrees, might be promising. Second, supporting the aforementioned argument, this dissertation shows the important role of vocational education and training for the development of non-cognitive skills. Policy makers should keep this role in mind when implementing or restructuring education and training systems. Third, this dissertation shows the value of screening periods, a substantial part of dual vocational education and training programs for young workers, as these periods allow employers to screen for non-cognitive skills. Therefore, promoting such programs could facilitate the school-to-work transitions by allowing individuals to communicate their valuable non-cognitive skills to potential employers. Thus the results in this dissertation suggest that well-developed vocational education and training programs are beneficial for the development of non-cognitive skills and their revelation to future employers.

This dissertation shows that labor market-relevant non-cognitive skills develop even during adolescence and adulthood, and that their development is related to education. Moreover, firms use non-cognitive skills—not cognitive ability—as their major selection criterion when making job offers, which is one potential explanation for the labor market importance of non-cognitive skills. Taken together, the results of my analyses are empirical support for the thesis that character is malleable and trumps intellect.

Bibliography

Acemoglu, D., & Pischke, J.-S. (1998). Why Do Firms Train? Theory and Evidence. *Quarterly Journal of Economics*, *113*(1), 79–119.

Acemoglu, D., & Pischke, J.-S. (1999). Beyond Becker: Training in Imperfect Labour Markets. *Economic Journal*, *109*(453), 112–142.

Agan, A. Y. (2014). Disaggregating the Returns to College. *Princeton University Working Paper*.

Allison, P. D. (1990). Change Scores as Dependent Variables in Regression Analysis. *Sociological Methodology*, *20*, 93–114.

Almlund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. (2011). Personality Psychology and Economics. In E. A. Hanushek, S. J. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp. 1–181). Amsterdam: Elsevier.

Altonji, J. G., & Pierret, C. R. (2001). Employer Learning and Statistical Discrimination. *Quarterly Journal of Economics*, *116*(1), 313–350.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Lost in State Space: Are Preferences Stable? *International Economic Review*, *49*(3), 1091–1112.

Andreoni, J., & Sprenger, C. (2012). Risk Preferences Are Not Time Preferences. *American Economic Review*, *102*(7), 3357–3376.

Andreoni, J., & Sprenger, C. (2015). Risk Preferences Are Not Time Preferences: Reply.

American Economic Review, *105*(7), 2287–2293.

Arcidiacono, P., Aucejo, E., Maurel, A., & Ransom, T. (2016). College Attrition and the Dynamics of Information Revelation. *NBER Working Paper Series*, *22325*.

Backes-Gellner, U., Herz, H., Kosfeld, M., & Oswald, Y. (2018). Do Preferences and Biases predict Life Outcomes? Evidence from Education and Labor Market Entry Decisions. *Swiss Leading House of Economics of Education Working Paper Series*, *144*.

Baert, S., & Decuypere, L. (2014). Better Sexy Than Flexy? A Lab Experiment Assessing the Impact of Perceived Attractiveness and Personality Traits on Hiring Decisions. *Applied Economics Letters*, *21*(9), 597–601.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does High Self-Esteem Cause Better Performance, Interpersonal Success, Happiness, or Healthier Lifestyles? *Psychological science in the public interest*, *4*(1), 1–44.

Becker, A., Deckers, T., Dohmen, T., Falk, A., & Kosse, F. (2012). The Relationship Between Economic Preferences and Psychological Personality Measures. *Annual Review of Economics*, *4*(1), 453–478.

Belfield, C. R., & Bailey, T. (2011). The Benefits of Attending Community College: A Review of the Evidence. *Community College Review*, *39*(1), 46–68.

Bénabou, R., & Tirole, J. (2002). Self-Confidence and Personal Motivation. *Quarterly Journal of Economics*, *117*(3), 871–915.

Bessey, D. (2010). *Educational Investment of Youths: Empirical and Experimental Evidence.* Retrieved from the Catalogue of University Dissertations UZH (No. 006206640): Dissertation University of Zurich.

BFS. (2017). *Lehrvertragsaufloesung, Wiedereinstieg, Zertifikationsstatus: Resultate zur*

Dualen Beruflichen Grundbildung (EBA und EFZ). Neuchatel: Bundesamt fuer Statistik.

Biesma, R. G., Pavlova, M., van Merode, G. G., & Groot, W. (2007). Using Conjoint Analysis to Estimate Employers Preferences for Key Competencies of Master Level Dutch Graduates Entering the Public Health Field. *Economics of Education Review*, *26*(3), 375–386.

Blascovich, J., & Tomaka, J. (1991). Measures of Self-Esteem. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 115–160). San Diego, Ca: Academic Press.

Bolli, T., & Hof, S. (2014). The Impact of Apprenticeship Training on Personality Traits: An Instrumental Variable Approach. *KOF Working Papers*, *350*.

Bolli, T., & Renold, U. (2017). Comparative Advantages of School and Workplace Environment in Skill Acquisition. *Evidence-based HRM*, *5*(1), 6–29.

Borghans, L., Golsteyn, B. H. H., Heckman, J. J., & Humphries, J. E. (2016). What Grades and Achievement Tests Measure. *Proceedings of the National Academy of Sciences*, *113*(47), 13354–13359.

Bound, J., & Turner, S. (2011). Dropouts and Diplomas: The Divergence in Collegiate Outcomes. In E. A. Hanushek, S. J. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp. 573–613). Amsterdam: Elsevier.

Bowles, S., Gintis, H., & Osborne, M. (2001). The Determinants of Earnings: A Behavioral Approach. *Journal of Economic Literature*, *39*(4), 1137–1176.

Brañas-Garza, P., Kujal, P., & Lenkei, B. (2015). Cognitive Reflection Test: Whom, How and When. *MPRA Working Paper*, *68049*.

Burks, S. V., Lewis, C., Kivi, P. A., Wiener, A., Anderson, J. E., Götte, L., . . . Rustichini,

A. (2015). Cognitive Skills, Personality, and Economic Preferences in Collegiate Success. *Journal of Economic Behavior & Organization*, *115*, 30–44.

CEDEFOP. (2014). Piloting a European Employer Survey on Skill Needs. *European Center for the Development of Vocational Training Research Papers*, *36*.

Cobb-Clark, D. A., & Schurer, S. (2012). The Stability of Big-Five Personality Traits. *Economics Letters*, *115*(1), 11–15.

Costa, P. T., & McCrae, R. R. (1994). Stability and Change in Personality from Adolescence through Adulthood. In C. F. Halverson, G. A. Kohnstamm, & R. Martin (Eds.), *The developing strucutre of temperament and personaility for infancy to adulthood* (pp. 139–150). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cunha, F., & Heckman, J. J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, *43*(4), 738–782.

Cunha, F., Heckman, J. J., Lochner, L. J., & Masterov, D. V. (2006). Interpreting the Evidence of Life Cycle Skill Formation. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (pp. 697–812). Amsterdam: Elsevier.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, *78*(3), 883–931.

Dahmann, S., & Anger, S. (2014). The Impact of Education on Personality: Evidence from a German High School Reform. *IZA Discussion Paper*, *8139*.

De Araujo, P., & Lagos, S. (2013). Self-Esteem, Education, and Wages Revisited. *Journal of Economic Psychology*, *34*, 120–132.

Deming, D. (2017). The Growing Importance of Social Skills in the Labor Market. *Quarterly Journal of Economics*, *132*(4), 1593-1640.

Deming, D., & Kahn, L. (2018). Skill Requirements Across Firms and Labor Markets: Evidence From Job Postings for Professionals. *Journal of Labor Economics*, *36*(1), 337–369.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review*, *100*(3), 1238–1260.

Donnellan, M. B., Hill, P. L., & Roberts, B. W. (2015). Personality Development Across the Life Span: Current Findings and Future Directions. In M. Mikulincer, P. R. Shaver, L. M. Cooper, & R. J. Larsen (Eds.), *Apa handbook of personality and social psychology.* Washington, D.C.: American Psychological Association.

Drago, F. (2011). Self-Esteem and Earnings. *Journal of Economic Psychology*, *32*(3), 480–488.

Duckworth, A. L. (2016). *Grit: The Power of Passion and Preserverance.* New York: Scribner.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and Passion for Long-Term Goals. *Journal of Personality and Social Psychology*, *92*(6), 1087–1101.

Duckworth, A. L., & Quinn, P. D. (2009). Development and Validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment*, *91*(2), 166–174.

Elkins, R., Kassenboehmer, S. C., & Schurer, S. (2016). The Malleability of Personality Traits in Adolesence. *Life Course Center Working Paper Series*, *2016-20*.

Eskreis-Winkler, L., Shulman, E. P., Beal, S. A., & Duckworth, A. L. (2014). The Grit Effect: Predicting Retention in the Military, the Workplace, School and Marriage. *Frontiers in Psychology*, *5*, 1–12.

Euwals, R., & Winkelmann, R. (2004). Training Intensity and First Labor Market Outcomes of Apprenticeship Graduates. *International Journal of Manpower*, *25*(5), 447–462.

Franz, W., & Zimmermann, V. (2002). The Transition From Apprenticeship Training to Work. *International Journal of Manpower*, *23*(5), 411–425.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42.

FSO. (2016). Newsletter 2/2016. *Federal Statistical Office*.

Gerlitz, J.-Y., & Schupp, J. (2005). Zur Erhebung der Big-Five-Basierten Persönlichkeitsmerkmale im SOEP. *DIW Research Notes*, *2005-4*.

Gibbons, R., & Katz, L. F. (1991). Layoffs and Lemons. *Journal of Labor Economics*, *9*(4), 351–380.

Goldsmith, A. H., Veum, J. R., & Darity, W., Jr. (1996). The Impact of Labor Force History on Self-Esteem and Its Component Parts, Anxiety, Alienation and Depression. *Journal of Economic Psychology*, *17*(2), 183–220.

Goldsmith, A. H., Veum, J. R., & Darity, W., Jr. (1997). The Impact of Psychological and Human Capital on Wages. *Economic Inquiry*, *35*(4), 815–829.

Hamilton, B. H., Papageorge, N. W., & Pande, N. (2014). The Right Stuff? Personality and Entrepreneurship. *SSRN Working Papers*, *2438944*.

Hanushek, E. A. (2002). Publicly Provided Education. In A. J. Auerbach, R. Chetty, M. Feldstein, & E. Saez (Eds.), *Handbook of public economics* (pp. 2045–2141). Amsterdam: Elsevier.

Harhoff, D., & Kane, T. J. (1997). Is the German Apprenticeship System a Panacea for the U.S. Labor Market? *Journal of Population Economics*, *10*(2), 171–196–196.

Heckman, J. J. (1995). Lessons From the Bell Curve. *Journal of Political Economy*, *103*(5), 1091–1120.

Heckman, J. J., Humphries, J. E., & Kautz, T. (Eds.). (2014). *The Myth of Achievement Tests: The GED and the Role of Character in America.* Chicago: The University of Chicago Press.

Heckman, J. J., Humphries, J. E., Urzua, S., & Veramendi, G. (2011). The Effects of Educational Choices on Labor Market, Health, and Social Outcomes. *Human Capital and Economic Opportunity Working Group, Working Papers*, *2011-002*.

Heckman, J. J., Humphries, J. E., Veramendi, G., & Urzua, S. (2014). Education, Health and Wages. *NBER Working Paper Series*, *19971*.

Heckman, J. J., & Kautz, T. (2012). Hard Evidence on Soft Skills. *Labour Economics*, *19*(4), 451–464.

Heckman, J. J., & Mosso, S. (2014). The Economics of Human Development and Social Mobility. *Annual Review of Economics*, *6*(1), 689–733.

Heckman, J. J., Pinto, R., & Savelyev, P. (2013). Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. *American Economic Review*, *103*(6), 2052–2086.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, *24*(3), 411–482.

Heineck, G., & Anger, S. (2010). The Returns to Cognitive Abilities and Personality Traits in Germany. *Labour Economics*, *17*(3), 535–546.

Herrnstein, R. J., & Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.

Himmler, O., & Koenig, T. (2012). Self-Evaluations and Performance: Evidence From Adolescence. *Diskussionspapiere der Wirtschaftswissenschaftlichen Fakultät, Universität Hannover*, *502*.

Hoeschler, P., & Backes-Gellner, U. (2014). Shooting for the Stars and Failing: College Dropout and Self-Esteem. *Swiss Leading House of Economics of Education Working Paper Series*, *100*.

Hoeschler, P., & Backes-Gellner, U. (2017). The Relative Importance of Personal Characteristics for the Hiring of Young Workers. *Swiss Leading House of Economics of Education Working Paper Series*, *142*.

Hoeschler, P., Balestra, S., & Backes-Gellner, U. (2017). Development of Non-Cognitive Skills in Adolescence. *Swiss Leading House of Economics of Education Working Paper Series*, *138*.

Hoeschler, P., Balestra, S., & Backes-Gellner, U. (2018). Development of Non-Cognitive Skills in Adolescence. *Economics Letters*, *163*, 40-45.

Hoffman, N., & Schwartz, R. (2015). *Gold Standard: The Swiss Vocational Education and Training System*. Washington, DC: National Center on Education and the Economy.

Humburg, M., & van der Velden, R. (2015). Skills and the Graduate Recruitment Process: Evidence From Two Discrete Choice Experiments. *Economics of Education Review*, *49*, 24–41.

Humphries, J. E., & Kosse, F. (2017). On the Interpretation of Non-Cognitive Skills: What Is Being Measured and Why It Matters . *Journal of Economic Behavior & Organization*, *136*, 174–185.

Jacobson, N. S., & Truax, P. (1991). Clinical Significance: A Statistical Approach to Defining Meaningful Change in Psychotherapy Research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19.

Jaeger, D. A., & Page, M. E. (1996). Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education. *Review of Economics and Statistics*, *78*(4), 733–740.

Jepsen, C., Troske, K., & Coomes, P. (2014). The Labor-Market Returns to Community College Degrees, Diplomas, and Certificates. *Journal of Labor Economics*, *32*(1), 95–121.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm Shift to the Integrative Big Five Trait Taxonomy. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). New York, NY: Guilford Press.

Kane, T. J., & Rouse, C. E. (1995). Labor-Market Returns to Two- and Four-Year College. *American Economic Review*, *85*(3), 600–614.

Kaufman, A. S., & Lichtenberger, E. O. (1998). Intellectual Assessment. In A. S. Bellack & M. Hersen (Eds.), *Comprehensive clinical psychology* (pp. 187–238). Oxford: Pergamon.

Kautz, T., Heckman, J. J., Diris, R., ter Weel, B., & Borghans, L. (2014). Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success. *NBER Working Papers*, *20749*.

Lazear, E. P. (1986). Raids and Offer Matching. *Research in Labor Economics*, *8*(A), 577–601.

Light, A., & Strayer, W. (2000). Determinants of College Completion: School Quality or Student Ability? *Journal of Human Resources*, *35*(2), 299–332.

Liu, V. Y. T., Belfield, C. R., & Trimble, M. J. (2015). The Medium-Term Labor Market

Returns to Community College Awards: Evidence From North Carolina. *Economics of Education Review*, *44*, 42–55.

Manning, A. (2011). Imperfect Competition in the Labor Market. In D. Card & O. Ashenfelter (Eds.), *Handbook of labor economics* (pp. 973–1041). Amsterdam: Elsevier.

Manski, C. F. (1992). Parental Income and College Oportunity. *Democratic Study Center, Education Resource Information Center Report ED349919*.

Meier, S., & Sprenger, C. D. (2015). Temporal Stability of Time Preferences. *Review of Economics and Statistics*, *97*(2), 273–286.

Mohrenweiser, J., Wydra-Somaggio, G., & Zwick, T. (2017). Soft Skills as a Source for Information Advantages of Training Employers. *Swiss Leading House of Economics of Education Working Paper Series*, *121*.

Muehlemann, S., Braendli, R., & Wolter, S. C. (2013). Invest in the Best or Compensate the Weak? *Evidence-based HRM*, *1*(1), 80–95.

Muehlemann, S., Pfeifer, H., Walden, G., Wenzelmann, F., & Wolter, S. C. (2010). The Financing of Apprenticeship Training in the Light of Labor Market Regulations. *Labour Economics*, *17*(5), 799–809.

Mueller, B., & Schweri, J. (2015). How Specific Is Apprenticeship Training? Evidence From Inter-Firm and Occupational Mobility After Graduation. *Oxford Economic Papers*, *67*(4), 1057–1077.

Mueller, G., & Plug, E. (2006). Estimating the Effect of Personality on Male and Female Earnings. *ILR Review*, *60*(1).

Murnane, R. J., Willett, J. B., Braatz, M. J., & Duhaldeborde, Y. (2001). Do Different Dimensions of Male High School Students' Skills Predict Labor Market Success a Decade Later? Evidence From the NLSY. *Economics of Education Review*, *20*(4), 311–320.

Neal, D., & Rosen, S. (2000). Theories of the Distribution of Earnings. In A. B. Atkinson & F. Bourguignon (Eds.), *Handbook of income distribution* (pp. 379–427). Amsterdam: Elsevier.

Nyhus, E. K., & Pons, E. (2005). The Effects of Personality on Earnings. *Journal of Economic Psychology*, *26*(3), 363–384.

Orth, U., Robins, R. W., & Widaman, K. F. (2012). Life-Span Development of Self-Esteem and Its Effects on Important Life Outcomes. *Journal of Personality and Social Psychology*, *102*(6), 1271–1288.

Oswald, Y. (2013). *Educational Decisions Under Uncertainity*. Retrieved from the Catalogue of University Dissertations UZH (No. 010043822): Dissertation University of Zurich.

Oswald, Y., & Backes-Gellner, U. (2014). Learning for a Bonus: How Financial Incentives Interact with Preferences. *Journal of Public Economics*, *118*, 52–61.

Park, J. H. (1996). Measuring Education Over Time: a Comparison of Old and New Measures of Eduction From the Current Population Survey. *Economics Letters*, *50*(3), 425–428.

Persico, N., Postlewaite, A., & Silverman, D. (2004). The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height. *Journal of Political Economy*, *112*(5), 1019–1053.

Prevoo, T., & ter Weel, B. (2015). The Effect of Family Disruption on Children's Personality Development: Evidence from British Longitudinal Data. *De Economist*, *163*(1), 61–93.

Protsch, P., & Solga, H. (2015). How Employers Use Signals of Cognitive and Noncognitive Skills at Labour Market Entry: Insights from Field Experiments. *European Sociological Review*, *31*(5), 521–532.

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2003). Work experiences and personality development in young adulthood. *Journal of Personality and Social Psychology*, *84*(3), 582–593.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2015). Patterns of Mean-Level Change in Personality Traits Across the Life Course: A Meta-Analysis of Longitudinal Studies. *Psychological Bulletin*, *132*(1), 1–25.

Roberts, B. W., Wood, D., & Caspi, A. (2010). The Development of Personality Traits in Adulthood. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 375–398). New York City: Guilford Press.

Robins, R. W., Trzesniewski, K. H., Tracy, J. L., Gosling, S. D., & Potter, J. (2002). Global Self-Esteem Across the Life Span. *Psychology and Aging*, *17*(3), 423–434.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global Self-Esteem and Specific Self-Esteem: Different Concepts, Different Outcomes. *American Sociological Review*, *60*(1), 141–156.

Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The Impact of Response Distortion on Preemployment Personality Testing and Hiring Decisions. *Journal of Applied Psychology*, *83*(4), 634–644.

Schönberg, U. (2007, October). Testing for Asymmetric Employer Learning. *Journal of Labor Economics*, *25*(4), 651–691.

Segal, C. (2012). Working When No One Is Watching: Motivation, Test Scores, and

Economic Success. *Management Science*, *58*(8), 1438–1457.

SERI. (2014). Vocational and Professional Education and Training in Switzerland. *State Secretariat for Education, Research and Innovation*.

SERI. (2017). Evaluation EBA II. *State Secretariat for Education, Research and Innovation*.

Stinebrickner, R., & Stinebrickner, T. (2008). The Effect of Credit Constraints on the College Drop-Out Decision: A Direct Approach Using a New Panel Study. *American Economic Review*, *98*(5), 2163–2184.

Stratton, L. S., O'Toole, D. M., & Wetzel, J. N. (2008). A Multinomial Logit Model of College Stopout and Dropout Behavior. *Economics of Education Review*, *27*(3), 319–331.

Stratton, L. S., & Wetzel, J. N. (2013). Are Students Dropping Out or Simply Dragging Out the College Experience? Persistence at the Six-Year Mark. *The B.E. Journal of Economic Analysis & Policy*, *13*(2).

Strupler, M., & Wolter, S. C. (2012). *Die duale Lehre: eine Erfolgsgeschichte-auch für Betriebe*. Zurich: Ruegger Verlag.

Teijeiro, M., Rungo, P., & Freire, M. J. (2013). Graduate Competencies and Employability: the Impact of Matching Firms' Needs and Personal Attainments. *Economics of Education Review*, *34*, 286–295.

Todd, P. E., & Wolpin, K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *Economic Journal*, *113*(485), F3–F33.

Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2003). Stability of Self-Esteem Across the Life Span. *Journal of Personality and Social Psychology*, *84*(1), 205–220.

von Aster, M., Neubauer, A., & Horn, R. (2009). *Wechsler-Intelligenztest für Erwachsene*. Frankfurt: Pearson.

von Wachter, T., & Bender, S. (2006). In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers' Careers. *American Economic Review*, *96*(5), 1679–1705.

Waddell, G. R. (2006). Labor-Market Consequences of Poor Attitude and Low Self-Esteem in Youth. *Economic Inquiry*, *44*(1), 69–97.

Waldman, M. (1990). Up-or-Out Contracts: A Signaling Perspective. *Journal of Labor Economics*, *8*(2), 230–250.

Wolpin, K. I. (2005). Education Data in the NLSY79. *Monthly Labor Review*, *128*, 15–20.

Wolter, S. C., & Ryan, P. (2011). Apprenticeship. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (pp. 521–576). Amsterdam: Elsevier.

Zeidenberg, M., Scott, M., & Belfield, C. (2015). What About the Non-Completers? The Labor Market Returns to Progress in Community College. *Economics of Education Review*, *49*, 142–156.

Additional Material for Chapter 3 and Chapter 4

Questionnaire of Leading House Apprenticeship Panel wave five (2014/2015)

Universität Zürich[UZH]

Vorname: ______________________________ Name: ______________________________

Bitte lesen Sie jede Aussage genau durch und wählen Sie als Antwort die Kategorie, die Ihrer Sichtweise am besten entspricht.

A. Übertritt in den Arbeitsmarkt

Bitte *kreuzen* Sie Zutreffendes an und *ergänzen* Sie das Textfeld.

1. Im Jahr 2009 haben Sie die Ausbildung zur Kauffrau/zum Kaufmann begonnen. Haben Sie diese Ausbildung abgeschlossen?

☐ Ja, im ☐ Sommer 2012.
☐ Sommer 2013.
☐ Sommer 2014.
☐ Nein, ich mache zurzeit ____.

2. Mit welcher Durchschnittsnote haben Sie Ihre Berufsausbildung abgeschlossen? Endnote: __________.

3. Hatten Sie direkt nach Abschluss Ihrer Lehre ein konkretes Übernahmeangebot von Ihrem Ausbildungsbetrieb?

☐ **Nein.** *(Weiter mit Frage 5)*
☐ **Ja. Der angebotene feste Monatslohn lag bei SFr. ... (Brutto)**

☐ unter 2'500	☐ 2'500 +	☐ 3'000 +	☐ 3'500 +	☐ 4'000 +	☐ 4'500 +
☐ 5'000 +	☐ 5'500 +	☐ 6'000 +	☐ 6'500 +	☐ 7'000 +	☐ über 7'500

Beinhaltete das Übernahmeangebot mehr als 12 Monatslöhne?

☐ Ja, 13 Monatslöhne.
☐ Ja, 14 Monatslöhne.
☐ Nein.

4. Haben Sie dieses Angebot angenommen?

☐ Ja.
☐ Nein, weil...
(Sie dürfen auch mehrere Antworten ankreuzen.)

☐ ich ein höheres Gehaltsangebot von einer anderen Firma hatte.
☐ ich eine interessantere Tätigkeit von einer anderen Firma angeboten bekam.
☐ mich eine andere Firma mehr interessierte.
☐ ich meinen Wohnort nach der Ausbildung ändern wollte.
☐ ich Teilzeit arbeiten wollte, und das im Ausbildungsbetrieb nicht möglich war.
☐ sonstiger Grund, nämlich ____.

5. Wie viele Monate haben Sie für die erste Stelle nach Ihrem Lehrabschluss gesucht?
Wie viele Bewerbungen haben Sie in diesem Zeitraum ungefähr verschickt?
Falls Sie noch immer auf der Suche sind, geben Sie bitte die Anzahl zum jetzigen Zeitpunkt an.

Anzahl Suchmonate: _________.

Anzahl verschickte Bewerbungen: _________.

Bitte weiter auf der Rückseite.

- 1 -

6. Sind Sie vom Lehrabschluss ohne Pause in eine Erwerbstätigkeit übergegangen?

☐ Ja.
☐ Nein. Nach dem Lehrabschluss war ich ohne Beschäftigung für _______ Monate, weil...

☐ ich diese Zeit für Ferien genutzt habe.
☐ ich diese Zeit für einen Sprachkurs genutzt habe.
☐ ich diese Zeit für einen Militär-, Zivilschutz- oder Zivildiensteinsatz genutzt habe.
☐ ich diese Zeit mit anderen Aktivitäten überbrückt habe, nämlich mit___________________________________.

7. Sind Sie noch beim gleichen Betrieb wie in der ersten Stelle nach Ihrem Lehrabschluss?

☐ Ja. *(Weiter mit Frage 9 im Abschnitt B)*
☐ Nein.

8. Bildeten alle Betriebe, bei denen Sie nach Abschluss Ihrer Lehre arbeiteten, Lehrlinge aus?

☐ Ja, alle.
☐ Nein, nicht alle.

B. Aktuelle Erwerbstätigkeit

Falls Sie mehr als eine Erwerbstätigkeit ausüben, beantworten Sie die Fragen bitte mit Fokus auf Ihre Haupterwerbstätigkeit.

9. Zurzeit bin ich...

☐ angestellt auf einer regulären Stelle.
☐ angestellt auf einem Praktikum.
☐ selbstständig.
☐ nicht erwerbstätig, aber...
(Sie dürfen auch mehrere Antworten ankreuzen.)

☐ ich bin aktiv auf der Suche nach einer Erwerbstätigkeit. *(Weiter mit Frage 19 im Abschnitt C)*
☐ ich bin in einer Vollzeitausbildung. *(Weiter mit Frage 19 im Abschnitt C)*
☐ ich leiste einen Militär-, Zivilschutz- oder Zivildienst-Einsatz. *(Weiter mit Frage 19 im Abschnitt C)*
☐ ich nehme mir bewusst eine Auszeit. *(Weiter mit Frage 19 im Abschnitt C)*
☐ ich übe eine andere Tätigkeit aus, und zwar ___________________________. *(Weiter mit Frage 19 im Abschnitt C)*

10. Ich bin zurzeit…

☐ Vollzeit erwerbstätig (100%).
☐ Teilzeit erwerbstätig (_______%), weil...
(Sie dürfen auch mehrere Antworten ankreuzen.)

☐ ich nebenher noch für andere Arbeitgeber arbeite.
☐ ich mich selbständig machen möchte.
☐ ich eine weitere Aus- oder Weiterbildung absolviere.
☐ ich so besser meinen Freizeit- oder Sportaktivitäten nachgehen kann.
☐ ich einen anderen Grund für mein Teilzeitpensum habe**,** nämlich _________________________________.

11. Mein Vertrag ist …

☐ unbefristet. ☐ befristet.

12. Ich arbeite zur Zeit …

☐ im selben Beruf wie in der Lehre.
☐ in einem anderen/neuen Beruf , nämlich als____.

und zwar ☐ in der gleichen Branche wie in der Lehre.
☐ in einer anderen Branche, nämlich in__________________________.

- 2 -

Universität Zürich UZH

13. Mein aktueller fester Monatslohn beträgt SFr. ... (Brutto)

☐ unter 2'500 ☐ 2'500 + ☐ 3'000 + ☐ 3'500 + ☐ 4'000 + ☐ 4'500 +

☐ 5'000 + ☐ 5'500 + ☐ 6'000 + ☐ 6'500 + ☐ 7'000 + ☐ über 7'500

14. Erhalten Sie weitere Lohnbestandteile? (*Sie dürfen auch mehrere Antworten ankreuzen.)*

☐ Nein.
☐ Ja, einen 13. Monatslohn.
☐ Ja, einen 14. Monatslohn.
☐ Ja, einen variablen Bestandteil, und zwar in Form von:

- ☐ Überstundenvergütung, ca. SFr.___________ pro Monat (Brutto).
- ☐ Trinkgelder, ca. SFr.___________ pro Monat (Brutto).
- ☐ Bonuszahlungen abhängig vom Team- oder Unternehmenserfolg, ca. SFr. ___________ pro Monat (Brutto).
- ☐ Leistungszulage abhängig von der eigenen Leistung, z.B. Stücklohn, Provision, ca. SFr. _______ pro Monat (Brutto).

15. Meine (vertraglich festgelegten) Ferien sind ______ Tage pro Jahr.

16. Mein Arbeitsweg dauert normalerweise ca. ______ Minuten (von Tür zu Tür) und ich verwende dabei meistens:

☐ Auto.
☐ ÖV.
☐ Velo.
☐ zu Fuss.
☐ anderes, nämlich:__________________.

17. Wie viele Mitarbeiter hat der Betrieb, beim dem Sie zurzeit tätig sind?

☐ 1 – 9
☐ 10 – 49
☐ 50 – 99
☐ 100 – 499
☐ 500 und mehr
☐ Nicht zutreffend, ich habe eine Einzelfirma gegründet.

18. Bildet der Betrieb, bei dem Sie zurzeit tätig sind, heute Lehrlinge aus?

☐ Ja.
☐ Nein.

C. Weiterbildung und berufliche Zukunft

19. Wie zufrieden sind Sie aus heutiger Perspektive mit Ihrer Ausbildung?

	Trifft gar nicht zu 1	2	3	4	**Trifft voll zu** 5
- Ich bin im Grossen und Ganzen sehr zufrieden.	☐	☐	☐	☐	☐
- Ich würde heute nochmals genau dieselbe Lehre wählen.	☐	☐	☐	☐	☐
- Meine Lehre hat mich optimal auf mein Berufsleben vorbereitet.	☐	☐	☐	☐	☐
- Andere Berufe würden mich heute mehr interessieren.	☐	☐	☐	☐	☐
- Ich habe mir von meiner Lehre höhere Lohn- oder Karrierechancen erwartet.	☐	☐	☐	☐	☐

Bitte weiter auf der Rückseite.

- 3 -

20. Machen Sie oder planen Sie zurzeit eine Weiterbildung oder weitere Ausbildung?

- Zweite Lehre	☐ Nein. ☐ Ja, geplant. ☐ Ja, begonnen. ☐ Ja, abgeschlossen. ☐ Weiss noch nicht.
- Berufs- oder Fachmaturität, Gymnasiale Maturität	☐ Nein. ☐ Ja, geplant. ☐ Ja, begonnen. ☐ Ja, abgeschlossen. ☐ Weiss noch nicht.
- Höhere Berufsbildung mit eidg. Fachausweis/Diplom/Meisterdiplom	☐ Nein. ☐ Ja, geplant. ☐ Ja, begonnen. ☐ Ja, abgeschlossen. ☐ Weiss noch nicht.
- Höhere Fachschule	☐ Nein. ☐ Ja, geplant. ☐ Ja, begonnen. ☐ Ja, abgeschlossen. ☐ Weiss noch nicht.
- Fachhochschule	☐ Nein. ☐ Ja, geplant. ☐ Ja, begonnen. ☐ Ja, abgeschlossen. ☐ Weiss noch nicht.
- Sonstiges, nämlich: ________________________	☐ Nein. ☐ Ja, geplant. ☐ Ja, begonnen. ☐ Ja, abgeschlossen. ☐ Weiss noch nicht.

21. Haben Sie sich schon einmal Gedanken gemacht, später eine eigene Firma zu gründen, oder selbständig erwerbstätig zu sein?

☐ Ja. ☐ Nein. *(Weiter mit Frage 23 im Abschnitt D)*

22. Wie konkret ist Ihre Idee für eine eigene Firma/selbständige Erwerbstätigkeit?

☐ Ich habe keine konkrete Idee für eine eigene Firma.
☐ Ich habe bereits eine konkrete Idee für eine eigene Firma, aber noch keine Schritte unternommen.
☐ Ich habe bereits eine konkrete Idee für eine eigene Firma und bereits erste Schritte unternommen.
☐ Ich habe bereits einen Geschäftsplan konkretisiert und/oder mit Kapitalgebern verhandelt.
☐ Ich besitze bereits eine eigene Firma, und beschäftige folgende Anzahl an Mitarbeitern:_____________.

D. Aktuelle persönliche Situation

23. Wie ist ihr Familienstand heute?

☐ Ledig.
☐ Verheiratet, mit Ehepartner zusammenlebend.
☐ In eingetragener Partnerschaft zusammenlebend.
☐ Verheiratet, dauernd getrennt lebend.
☐ In eingetragener Partnerschaft dauernd getrennt lebend.
☐ Geschieden/ Aufgelöste eingetragene Partnerschaft.
☐ Verwitwet/ Lebenspartner/-in aus eingetragener Partnerschaft verstorben.

24. Wie wohnen Sie heute?

☐ Alleine. ☐ In einer Wohngemeinschaft.
☐ Mit meinen Eltern. ☐ Mit meiner Partnerin / meinem Partner.

25. Wie viel Zeit verwenden Sie für Ihre Hobbies? ______ Stunden pro Woche.

26. Wie schätzen Sie sich persönlich ein:

	Sehr ungeduldig										**Sehr geduldig**
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Sind Sie im Allgemeinen ein Mensch, der ungeduldig ist, oder der immer sehr viel Geduld aufbringt?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	Gar nicht risikobereit										**Sehr risikobereit**
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Sind Sie im Allgemeinen ein risikobereiter Mensch oder versuchen Sie, Risiken zu vermeiden?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

- 4 -

Universität Zürich[UZH]

**27. Man kann sich in verschiedenen Bereichen ja unterschiedlich verhalten.
Wie würden Sie Ihre Risikobereitschaft in Bezug auf die folgenden Bereiche einschätzen?**

	Gar nicht risikobereit 0	1	2	3	4	5	6	7	8	9	Sehr risikobereit 10
- beim Autofahren?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
- bei Geldanlagen?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
- bei Freizeit und Sport?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
- bei Ihrer beruflichen Karriere?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
- bei Ihrer Gesundheit?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
- beim Vertrauen in fremde Menschen?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

28. Ich bin jemand der...

	Trifft gar nicht zu 1	2	3	4	5	6	Trifft voll zu 7
- gründlich arbeitet.	☐	☐	☐	☐	☐	☐	☐
- kommunikativ, gesprächig ist.	☐	☐	☐	☐	☐	☐	☐
- manchmal etwas grob zu anderen ist.	☐	☐	☐	☐	☐	☐	☐
- originell ist, neue Ideen einbringt.	☐	☐	☐	☐	☐	☐	☐
- sich oft Sorgen macht.	☐	☐	☐	☐	☐	☐	☐
- verzeihen kann.	☐	☐	☐	☐	☐	☐	☐
- eher faul ist.	☐	☐	☐	☐	☐	☐	☐
- aus sich heraus gehen kann, gesellig ist.	☐	☐	☐	☐	☐	☐	☐
- künstlerische Erfahrungen schätzt.	☐	☐	☐	☐	☐	☐	☐
- leicht nervös wird.	☐	☐	☐	☐	☐	☐	☐
- Arbeiten wirksam und effizient erledigt.	☐	☐	☐	☐	☐	☐	☐
- zurückhaltend ist.	☐	☐	☐	☐	☐	☐	☐
- rücksichtsvoll und freundlich mit anderen umgeht.	☐	☐	☐	☐	☐	☐	☐
- eine lebhafte Phantasie, Vorstellungen hat.	☐	☐	☐	☐	☐	☐	☐
- entspannt ist, mit Stress gut umgehen kann.	☐	☐	☐	☐	☐	☐	☐
- wissbegierig ist.	☐	☐	☐	☐	☐	☐	☐

29. Wie schätzen Sie sich ein?

	Trifft gar nicht zu 1	2	3	4	Trifft voll zu 5
- Neue Ideen und Projekte lenken mich manchmal von alten Ideen und Projekten ab.	☐	☐	☐	☐	☐
- Rückschläge entmutigen mich nicht.	☐	☐	☐	☐	☐
- Ich war kurzfristig von einer Idee oder einem Projekt besessen, habe aber später das Interesse daran verloren.	☐	☐	☐	☐	☐
- Ich arbeite hart.	☐	☐	☐	☐	☐
- Ich setze mir oft ein Ziel und beschliesse dann später, ein anderes Ziel zu verfolgen.	☐	☐	☐	☐	☐
- Ich habe Schwierigkeiten, mich auf Projekte zu konzentrieren, die länger als ein paar Monate bis zum Abschluss benötigen.	☐	☐	☐	☐	☐
- Ich bringe zu Ende, was auch immer ich angefangen habe.	☐	☐	☐	☐	☐
- Ich bin fleissig.	☐	☐	☐	☐	☐

30. Wenn Sie an Ihre heutige persönliche Situation denken, wie zutreffend sind die folgenden Aussagen?

	Trifft gar nicht zu 1	2	3	4	Trifft voll zu 5
Mit meinem Leben als Ganzes bin ich sehr zufrieden.	☐	☐	☐	☐	☐
Mit meiner finanziellen Situation bin ich im Grossen und Ganzen sehr zufrieden.	☐	☐	☐	☐	☐

Bitte weiter auf der Rückseite.

- 5 -

E. Verlosung iPad Mini / PlayStation 4

Wie angekündigt nehmen Sie wie immer an einer Verlosung teil. Dieses Mal können Sie zwischen zwei Preisen auswählen:

- ☐ Ich möchte ein Apple iPad Mini mit Retina Display, wenn ich gewinne.
- ☐ Ich möchte eine Sony PlayStation 4, wenn ich gewinne.

Damit wir Sie im Falle eines Gewinns kontaktieren können, geben Sie hier bitte Ihre vollständigen Kontaktdaten an:

Vorname, Name

Strasse — Hausnummer

Wohnort — PLZ

Handynummer

E-Mail

Unsere Studie wird voraussichtlich auch in einem Jahr wieder fortgeführt. Gerne werden wir Sie also zu diesem Zeitpunkt wieder kontaktieren und Sie werden dann wieder an einer Verlosung teilnehmen.

Wenn Sie uns noch Kommentare zukommen lassen möchten, dann freuen wir uns darüber sehr.
Bitte notieren Sie diese hier:

So, nun haben Sie es geschafft.
Für Ihre Mithilfe bedanken wir uns nochmals ganz herzlich und wünschen Ihnen für Ihre zukünftige berufliche Laufbahn viel Erfolg und alles Gute!

Falls Sie Fragen an uns haben, können Sie sich jederzeit bei uns melden.

Prof. Dr. Uschi Backes-Gellner und Peter Höschler
Lehrstuhl für Betriebswirtschaftslehre, insbes. empirische Methodik der Arbeitsbeziehungen und der Personalökonomik
Plattenstrasse 14
CH-8032 Zürich
Tel: 044 634 42 74
E-Mail: profbackesgellner@business.uzh.ch

Curriculum Vitae

Personal Details

Name:	Peter Höschler
Date of birth:	May 12, 1989
Nationality:	German

Education

09/2011-02/2018: Doctoral program at the University of Zurich, Graduate School of Business.

09/2011-02/2018: Course Program for Doctoral Students, Swiss Leading House on Economics of Education, Firm Behavior, and Training Policies.

09/2015-03/2016: Visiting doctoral fellow at the University of Chicago.

09/2011-04/2014: Master of Science, University of Zurich, Management and Economics specialization.

10/2008-09/2011: Bachelor of Science, University of Würzburg, Economics and Business Management specialization.

08/2010-05/2011: Visiting Student at the University of Texas at Austin.

Professional Experience

09/2011-01/2019: Research and Teaching Assistant, University of Zurich, Chair for Empirical Research in Business, Industrial Relations and Human Resource Management.

10/2009-08/2011: Student Assistant, University of Würzburg, Chair for Human Resource Management and Organization.